IRE Transactions on INFORMATION THEORY

A Journal Devoted to the Theoretical and Experimental Aspects of Information Transmission, Processing and Utilization.

Volume IT-5 SEPTEMBER, 1959 Number 3
Published Quarterly

In This Issue

PUBLISHED BY THE
Professional Group on Information Theory

IRE Professional Group on Information Theory

The Professional Group on Information Theory is an organization, with the framework of the IRE, of members with principal professional interest in Information Theory. All members of the IRE are eligible for membership in the Group and will receive all Group publications upon payment of an annual fee of $3.00.

IRE TRANSACTIONS® on INFORMATION THEORY is published by the IRE for the Professional Group on Information Theory, at 1 East 79th Street, New York 21, N. Y. Responsibility for contents rests upon the authors and not upon the IRE, the Group, or its members. Price per copy: IRE-PGIT members, $1.55; IRE members, $2.35; nonmembers, $4.65.

INFORMATION THEORY

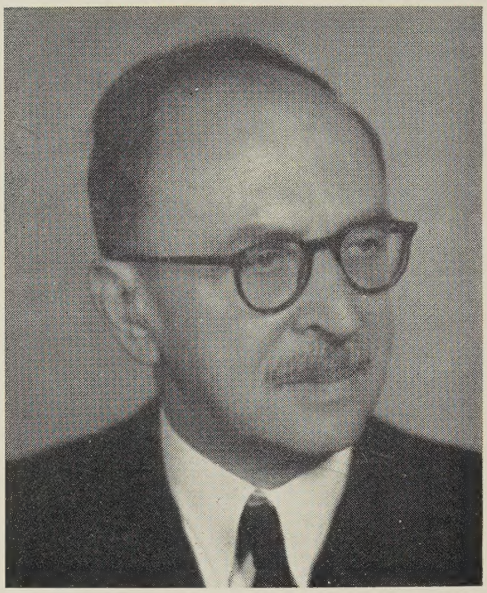
IRE Transactions on Information Theory

A Journal Devoted to the Theoretical and Experimental Aspects of Information Transmission, Processing and Utilization

Volume IT-5 | **September, 1959** | **Number 3**

Published Quarterly

TABLE OF CONTENTS



D. Gabor

D. Gabor, Professor of Applied Electron Physics in the Imperial College of Science and Technology, London, was born in Budapest, Hungary, and studied electrical engineering in Budapest and in Berlin, where he acquired the Degree of Dr.-Ing. in 1927. After seven years in a research laboratory of Siemens & Halske, Berlin, he worked from 1934 until 1948 in the research laboratory of the British Thomson-Houston Company, Rugby. He joined Imperial College in 1949, and was elected Fellow of the Royal Society in 1956. His principal scientific work relates to high speed oscillography, gas discharges, physical optics, communication theory and techniques, and all types of electronic devices. His present interests include plasma physics and "intelligent" machines.

Guest Editorial

IT is now a little over ten years since the appearance of Shannon's COMMUNICATION THEORY, and it is time to review the harvest it has brought in, scientific nd practical.

For one thing, INFORMATION THEORY (as it was soon lled) has now become mathematically respectable. Pure athematicians like Doob will no longer query "whether ie mathematical extensions of the author (Shannon) are ways strictly honorable." The work of McMillan, einstein, Khinchin and of Shannon himself has made NFORMATION THEORY rigorous—and almost unreadable to ngineers! It is curious to remember that thirty years rlier STATISTICAL MECHANICS had a similar adventure, ut with a difference. Outstanding mathematicians—Viener, Birkhoff, v. Neumann—could not accept the appy-go-lucky theories of the physicists Boltzmann and ibbs, and they put STATISTICAL MECHANICS, in particular ie Ergodic Theorem, on a rigorous mathematical basis. he only difference was that the physicists refused to ake any notice of this. By the time the mathematicians ad finished the ergodic theory of continuous processes, ie physicists had discovered quantum statistics, and efused to give physical significance to anything but what ie mathematician Rényi calls "entropies of zero order." hey never bothered to make Lebesgue-Borel measure ieory a part of the physical curriculum.

With so much effort of the best brains spent on founda-ons and on coding, it is not surprising that the practical chievements of COMMUNICATION THEORY fall a little ort of what could be expected ten years ago. It is not nfair to say that most progress has been made where ommunications were already near-perfect. We now have omplicated near-optimum codes which will operate with n error of 10^{-8}—10^{-9}, but the saving in bandwidth is ather less than 10 per cent compared with a much simpler rror-checking code operating under the same conditions. n the other hand, hardly any progress has been made in ie filtering of speech, or in the transmission of speech r pictures in reduced wavebands, where the potential aving is not measured in per cents but in orders of agnitude.

The lack of progress in filtering beyond the early work f Wiener and of Lee is due to a great extent, of course, o the enormous mathematical difficulties of a nonlinear lter theory. But it must be admitted that even if we had uch a theory, it would not be much good until we know hat it is that we want to filter out. The communication ystems which are most backward are those which ulti-nately serve the human receptor: the ear or the eye. Communication between machines can be (with rare xceptions) digitalized, and is adequately covered by resent efforts. But until we know more exactly what it s that a human ear picks out in a conversation with a oisy background, or what the eye picks out "at a glance" n a picture, no serious progress is possible, except per-aps by lucky inventions.

We still do not possess in speech perception anything comparable to the Maxwell-König theory of color perception, or Schrödinger's beautiful variant of it, based on the minimum perceptible interval in a three-dimensional non-euclidian metric. In view of Edwin Land's recent observations, I do not contend that these are solid foundations which will stand for all times, but I wish we had at least that much solid basis for a theory of speech perception. We can guess from the work of Walter Lawrence that speech-sound perception is six-dimensional. Six parameters (larynx frequency and intensity, hiss and the frequency of the first three formants) are sufficient for synthetizing very human-sounding speech. What is required is a mathematical theory of the selector operators for these parameters, and empirical work for establishing the *limen* or smallest perceptible step in this or an equivalent 6-dimensional space.

In picture transmission the situation is even more complicated. The experiments of the Gestalt psychologists have unearthed a confusing wealth of material relating to the perception of still pictures, and D. M. McKay's recent work has proved that time and motion make the chaos even worse. This may well be the muddle which, according to A. N. Whitehead, must precede a successful induction, but in the present case one feels inclined to ask: "Precede by how many hundred years?" It may well be that a device which removes noise from television pictures or allows transmitting them in a greatly reduced waveband cannot be much simpler than the human visual cortex. But to believe that no progress is possible before we have understood every optical illusion and have built it into a solid theoretical framework is dangerous defeatism. There is a gap of a million or so between the 20 bits per second which, the psychologists assure us, the human eye is capable of taking in, and what we offer it in a television picture. For my part, I believe that at least the first two orders of magnitude out of the six ought to be relatively easy going, and could be overcome on the basis of such elementary knowledge as that the eye fixes in the first line on contours, and samples what is in between (grass, hair, leaves, fabrics). This is a challenging problem for the *signal analyst*, who must first pioneer the field before statistical communication theory can start. To start at the other end, and begin collecting conditional probabilities of signals *before one knows what the signals really are* is one of the commonest errors of those who have only superficially digested information theory.

I wanted to dwell on the shortcomings of information theory rather than on its many positive achievements, not out of perversity, but because I believe that if more effort is directed into the No-Man's Land between raw sensory data and the distinguishable signals which are the starting point of the statistical theory, the second decade of INFORMATION THEORY will be as rich in practical improvements in communication techniques as the first was in intellectual clarifications.

—D. GABOR

Note on Unique Decipherability*

E. T. JAYNES†

Summary—We consider an alphabet of a letters, used under the restrictions: 1) messages uniquely decipherable into words by use of one of the letters as a space mark, and 2) words limited to a maximum length of L letters. Although imposing these constraints simultaneously may cause a large reduction in the channel capacity of the alphabet, neither by itself causes any reduction. Accordingly, in the absence of constraints other than 1), an inequality of McMillan pertaining to uniquely decipherable messages can be made to be an equality.

Defining "semi-optimal" transmission by the condition that the mean transmission time per word is minimized for a given entropy per word, we find the attainable rate of information transmission under semi-optimal conditions. Transmission at full channel capacity is a special case of semi-optimal transmission. Some generalizations and analogies to statistical mechanics are discussed.

Introduction

THE purpose of this note is twofold: 1) to give a result related to an inequality of McMillan pertaining to unique decipherability, and 2) to illustrate the close relationship between problems of information transmission and some of the elementary problems of statistical mechanics by use of notation borrowed from the latter field. In statistical mechanics we find that in principle all thermodynamic properties of a system are determined if we can evaluate its partition function Z; or better still, $\log Z$, in its dependence on the various constraints representing experimentally imposed conditions. Similarly, many problems of information transmission under constraints are, in principle, solved if we can evaluate an appropriate partition function. For the calculation of channel capacity, this is equivalent to the method described by Shannon. The same mathematical procedure also solves a wider class of problems, in which we find the transmission rate under what are termed "semi-optimal" conditions.

Channel Capacity Under Constraints

We have an alphabet of a symbols, each of which can be transmitted in unit time. Let l_i be the length (number of letters) of word w_i, and define the partition function[1]

$$Z(\lambda) = \sum_i 2^{-\lambda l_i} \tag{1}$$

where the sum is over all words in our vocabulary. A given vocabulary (*i.e.*, a specific set of possible words) may be regarded as defining a channel. By a theorem o Shannon,[2] the capacity of this channel is the larges (actually the only) real root of $Z(\lambda) = 1$.

The notion of a "word" is meaningful only if ther exists some rule by which a sequence of letters can b uniquely deciphered into words. If no such rule exists then effectively each letter is a word. The partition func tion then reduces to $Z(\lambda) = a2^{-\lambda}$, and Shannon's theorem gives the well-known channel capacity (all logarithms are to the base 2)

$$C = \log a \text{ bits/symbol.} \tag{2}$$

A necessary and sufficient condition for unique de cipherability into words (UD) is given by Sardinas and Patterson.[3] McMillan[4] has given two inequalities implied by UD, which had been noted before[5,6] under more re strictive conditions.[7] The first, which perhaps deserves to be called the fundamental inequality of noiseless coding theory, is in our notation

$$Z(\log a) \leq 1. \tag{3}$$

Although, as several authors have shown,[4-8] this in equality can be derived without any reference to infor mation theory, the concepts introduced by Shannon give it a simple intuitive meaning.

Any UD coding method is a system of constraints which in some way restricts our freedom in choosing the successive letters of a message, and defines a particular channel. Eq. (3) expresses the fact that imposing these con straints can never lead to a channel with greater capacity than the value (2), which corresponds to complete freedom of choice. Thus we conjecture that (3) will be fundamental not only for UD, but also for coding systems designed for any other objective. In general, a constraint will reduce channel capacity, and a reasonable measure of the ef ficiency of a code is the amount of this decrease.

* Manuscript received by the PGIT, Jan. 21, 1959. This research was supported by the USAF under Contract No. AF 49 (638)-342 monitored by the AF Office of Sci. Res. of the Air Res. and Dev. Command.

† Microwave Lab. and Dept. of Phys., Stanford University, Stanford, Calif.

[1] E. Schrödinger, "Statistical Thermodynamics," Cambridge University Press, Cambridge, Eng., chap. II; 1948.

[2] C. E. Shannon, "The Mathematical Theory of Communica tion," University of Illinois Press, Urbana, Illinois, p. 8; 1949.

[3] A. A. Sardinas and G. W. Patterson, "A necessary and sufficien condition for unique decomposition of encoded messages," IRE Convention Record, pt. 8, pp. 104–108; 1953.

[4] B. McMillan, "Two inequalities implied by unique decipher ability," IRE Trans. on Information Theory, vol. IT-2, pp 115–116; December, 1956.

[5] L. G. Kraft, "A device for quantizing, grouping and coding amplitude modulated pulses," S. M. Thesis, Dept. Elec. Eng. M.I.T., Cambridge, Mass.; 1949.

[6] B. Mandelbrot, "On recurrent noise limiting coding," *Proc Symp. on Information Networks*, Polytechnic Inst. of Bklyn; New York, N. Y.; 1955.

[7] Given also in A. Feinstein, "Foundations of Information Theory," McGraw-Hill, New York, N. Y., pp. 17–23; 1958.

[8] M. P. Schützenberger and R. S. Marcus, "Full decodable code-word sets," IRE Trans. on Information Theory, vol. IT-5 pp. 12–15; March, 1959.

McMillan[4] considers also a strong sufficient condition r UD, called irreducibility, and shows that for fixed ord lengths the strong constraint of irreducibility does ot reduce channel capacity below that set by the general nstraint of UD.

Irreducibility ensures UD only if the entire message is vailable. In applications to communication systems and genetics,[9] the most common transmission defect is the ne wherein certain parts of the message are simply lost. lthough the probability that this would destroy UD for ne entire balance of the message is usually very small,[10] ne is led to ask for a stronger condition of UD "in the nall." Without attempting a precise definition of rm, we use it in the rough sense that any reasonably ng fragment of a message will still be uniquely decipher-ble, except for possible end-effects.

Golomb, Gordon, and Welch[9] consider channels with xed word length k, and a structure constraint much ronger than UD, which ensures UD in the small. Among neir results, they show (theorems 6 and 7) that for given in the limit of large alphabets even their strong con-raint does not reduce capacity below the value (2).

Suppose we achieve UD by the usual method of choos-g one of the letters, which we call the "space," and sing it only as the terminating letter of each word; call nis "spacing." Spacing is a stronger constraint than IcMillan's irreducibility, but weaker than that of olomb, Gordon, and Welch (although it still accomplishes ne aim of UD in the small). We wish to find how much ne channel capacity is reduced by spacing, and to show nat this reduction is in fact zero, independently of the size f the alphabet, if spacing is the only constraint.

Due to the spacing constraint, the maximum number of ifferent words of length l is not a^l, but only

$$(a-1)^{l-1}.$$

Evidently, any failure to include all of the short words n our vocabulary will have a further adverse effect on nannel capacity, in addition to that imposed by spacing. Therefore we use all possible words of total length $l \leq L$, nd the partition function (1) becomes

$$Z(\lambda) = \sum_{l=1}^{L} (a-1)^{l-1} 2^{-\lambda l} = \frac{1-(a-1)^L 2^{-\lambda L}}{2^\lambda - a + 1}. \tag{4}$$

Noting that for all real λ, $Z(\lambda)$ is a decreasing function, nd that $Z(\lambda) \to L/(a-1)$ as $\lambda \to \log(a-1)$, it follows nat if $L = (a-1)$, the exact channel capacity is $C = \log$ $-1)$. If $L > (a-1)$, we find from (4) the inequalities

$$\log(a-1) < C \leq \log a. \tag{5}$$

The latter inequality in (5) is identical to (3), and it goes to an equality in the limit $L \to \infty$. But since $\log a$ is just the channel capacity we would have with an alphabet of a letters without any constraints, our assertion is proved. Stated differently, since (3) must hold for any method of achieving UD, *in the absence of other constraints, no method of achieving UD can be more efficient, as measured by channel capacity, than spacing.*

If $L < (a-1)$, then (4) leads to the inequalities

$$\left(1 - \frac{1}{L}\right) \leq \frac{C}{\log(a-1)} < 1, \tag{6}$$

the equality sign holding in the limit $a \to \infty$.

Numerical values of C obtained from (4) are given in Fig. 1. The trend for different a may seem disconcerting; one might argue that we are merely tying up one of the letters for special use, and so the loss in channel capacity could never exceed that due to removal of a single letter from the alphabet. However, the loss in capacity is in fact greatest for the large alphabets.

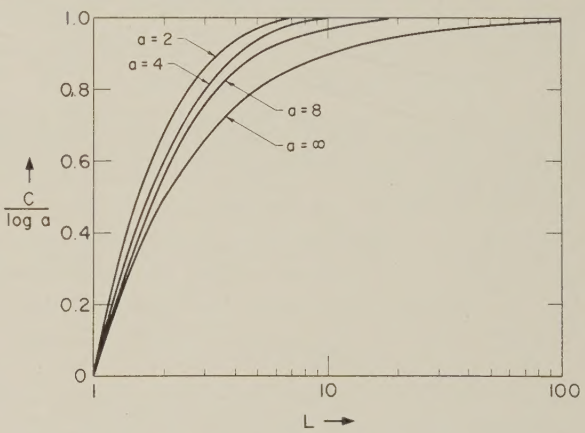


Fig. 1—Reduction in channel capacity due to restriction on maximum word length, for various alphabet sizes.

This situation may be understood as follows. Under no restrictions on word length, $L \to \infty$, we show in the next section that operation at full channel capacity requires a mean word length $\langle l \rangle = a$. The space then occurs with the same relative frequency, $(1/a)$, that it would have if it were not assigned any special function. This is the reason why UD, by itself, need not restrict transmission rate; the space is being used just as efficiently as any other symbol. However, when a becomes large, the effect of fixed L is to force $\langle l \rangle < a$. It is this tying up of channel time by too frequent repetition of the space which actually causes all the decrease in channel capacity, and explains the lower position, in Fig. 1, of the curves for large a.

Semi-Optimal Transmission

If the word w_i occurs with probability,

$$p_i = \frac{2^{-\lambda l_i}}{Z(\lambda)}, \tag{7}$$

the rate of transmission (entropy per word) is maximized

[9] S. W. Golomb, Basil Gordon, and L. R. Welch, "Comma-free odes," *Canadian Jour. Math.*, vol. 10, no. 2, pp. 202–209; 1958.
[10] M. P. Schützenberger, "On an application of semi-group ethods to some problems in coding," IRE Trans. on Information Theory, vol. IT-2, pp. 47–60; September, 1956.

for a given mean length of word; equivalently, the average transmission time per word is minimized for a given entropy per word. The average length and entropy per word, under these semi-optimal conditions, are given by[11]

$$\langle l \rangle = -\frac{\partial}{\partial \lambda} \log Z(\lambda) \tag{8}$$

$$S = \log Z(\lambda) + \lambda \langle l \rangle, \tag{9}$$

which are parametric equations connecting the quantities of interest. From them we can construct the "operating characteristic" of the channel, in which we plot the time rate of transmission,

$$H = \frac{S}{\langle l \rangle} \text{ bits/symbol}, \tag{10}$$

as a function of the average word length $\langle l \rangle$. A family of operating characteristics, computed from (4) in the case of no restriction on maximum word length, *i.e.*, from the partition function

$$Z(\lambda) = (2^{\lambda} - a + 1)^{-1}, \tag{11}$$

is given in Fig. 2 for various alphabet sizes. The range of attainable operating conditions consists of all points lying below the curve.

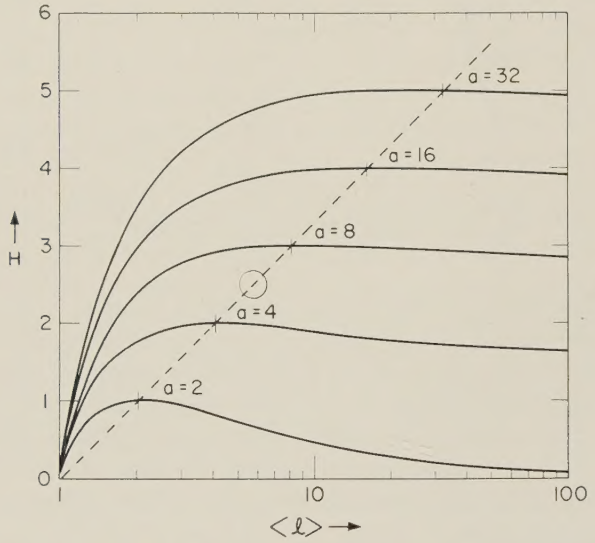


Fig. 2—Maximum attainable transmission rate as a function of average word length, for various alphabet sizes. For each curve, operation at full channel capacity occurs at the intersection with the dotted line.

The equation represented in Fig. 2 is found by eliminating λ from the above equations:

$$H = \log \langle l \rangle + \frac{\langle l \rangle - 1}{\langle l \rangle} \log \left[\frac{a - 1}{\langle l \rangle - 1} \right]. \tag{12}$$

[11] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, pp. 620–630, May 15, 1957; vol. 108, pp. 171–190, October 15, 1957.

It is in the limit $\langle l \rangle \to \infty$ that we are effectively removing one letter from the alphabet, and so $H \to \log (a - 1)$.

From (8), (9), and (10) the condition for maximum transmission rate is

$$\frac{dH}{d\langle l \rangle} = -\frac{\log Z}{\langle l \rangle^2} = 0, \tag{13}$$

or $\log Z(\lambda) = 0$. Under these conditions, we find $H = \lambda = C$, the channel capacity; thus (8) and (9) provide a simple alternative derivation of Shannon's rule for calculating channel capacity.[2]

In the case of the partition function (11), operation at full channel capacity occurs when $H = \lambda = C = \log a$; and (8) then gives $\langle l \rangle = a$, as previously noted.

The case $a = 32$ corresponds to the English language, if we consider the space and any five punctuation marks as included in the alphabet. The average word length in English is far less than 32 symbols, and varies with the source. One thousand consecutive words from Shannon's fundamental paper[2] had a mean length (including the space) of 5.9 symbols; while a similar analysis of James Michener's "Sayonara" gave a mean length of only 5.4. From Fig. 2, we find that because of Michener's tendency to use short words, unique decipherability by spacing costs him 0.3 bit per symbol in information content, while Shannon's loss was only 0.25.

The many additional constraints in English cause the actual transmission rate to fall considerably below the semi-optimal rate. Taking Shannon's estimate[12] of the redundancy of English as about 50 per cent, the actual operating region of English text would be given roughly by the circle in Fig. 2. From this we see that 1) only about 15 per cent of the redundancy is due to use of short words, and 2) the same rate of information transmission and the same mean word length to which we are accustomed could be achieved with an alphabet of only 6 symbols (5 letters and a space), if used at maximum efficiency.

Generalization

The above relations are easily generalized to the case where the transmission time is different for different symbols, and where we have other types of constraints.

[12] Shannon, *loc. cit.*, p. 26. See also C. E. Shannon, "Prediction and Entropy of Printed English," *Bell Sys. Tech J.*, vol. 30, pp. 50–64; January, 1951. Here the estimated redundancy is increased to about 75 per cent, from experiments in which human subjects attempted to restore missing parts of English text. However, the ability to do this may depend on semantic as well as purely statistical factors; and in any event the only properties which could be utilized for encoding efficiently into a smaller alphabet, are known frequencies. Estimates based on measured letter and word frequencies remain not much greater then 50 per cent, the value used above. Even these measurements suffer from fundamental ambiguities, some of which were pointed out by G. A. Barnard, "Statistical calculation of word entropies for four western languages," IRE Trans. on Information Theory, vol. IT-1, pp. 49–53; March, 1955. Fundamentally, of course, it is meaningless to say that there exists one and only one "true" redundancy for English text; one can speak only of the redundancy corresponding to certain specified statistical information.

or example, let the k'th letter of the alphabet have transission time t_k, and denote the space by the a'th letter. hen in (1) we interpret l_i as the total transmission time word w_i, and its evaluation is again elementary, with ie result

$$Z(\lambda) = \frac{2^{-\lambda t_a}}{1 - Q(\lambda)}, \tag{14}$$

here

$$Q(\lambda) \equiv \sum_{k=1}^{a-1} 2^{-\lambda t_k}.$$

rom this the channel capacity and transmission rate nder semi-optimal conditions may be found. Constraints the form that certain combinations of letters do not cur may lead to involved mathematical problems, but ot to any new difficulties of principle. Shannon's Theorem shows[2] that one common type of constraint is included we generalize the partition function to a "partition atrix," operation at full channel capacity then occurring hen the greatest eigenvalue of this matrix is unity.

Of course, the quantity t_k above need not be interpreted s a time. It can equally well stand for the "cost," as easured on any basis, of transmitting the k'th symbol. he theory then gives us the method of transmitting hich is most economical with respect to this cost assignent; semi-optimal transmission minimizes the average ost per word for a given entropy per word.

Relation to Statistical Mechanics

The partition function (11), with the number 2 replaced y e, is almost identical to the one arising in quantum atistical mechanics, describing a harmonic oscillator. he operating characteristic in Fig. 2, for the case $a = 2$, epresents nothing more than an unconventional way of lotting the Einstein specific heat function of a harmonic scillator.

Each of the soluble problems of statistical mechanics so provides the solution to a certain problem of infoation transmission under constraints, and the mathmatical analogy may be set up in other ways than the ne indicated here. In the above type of analogy, noted efore by Mandelbrot,[6] the mean word length correponds to the thermodynamic energy function, the paramter λ to the reciprocal temperature. Thus the transmission te H corresponds, not to the thermodynamic entropy nction, but to the ratio (entropy)/(energy).

It is interesting that such a fundamental notion as annel capacity has no thermodynamic analog. In ermodynamics the absolute value of the entropy has meaning; only entropy differences can be measured in xperiments. Consequently the condition that H is aximized, equivalent to the statement that the Helmholz ee energy function vanishes ($A \equiv E - TS = 0$), correponds to no condition which could be detected experientally.

Generalization of the above analysis to the case where the different words are no longer statistically independent is also straightforward, and corresponds to the transition in statistical mechanics from the Maxwell-Boltzman "molecular" viewpoint, to the Gibbsian "global" viewpoint.[1] We mention two examples of the correspondence which then exists.

The partition function of the linear Ising chain,[13] with an easy generalization, provides also an explicit solution to the problem of encoding a message into binary digits, in a way which is optimal from the standpoint of a person who knows the digram frequencies of the source, but has no other statistical information. The corresponding solution for trigrams would be of considerable interest in connection with the theory of ferromagnetism.

The two-dimensional Ising model of ferromagnetism,[13] the partition function of which was first obtained by Onsager, gives the solution to a problem in which a message in binary digits has strong correlations between adjacent symbols, and also between the n'th and the $(n + M)$'th, where M is a large fixed number. Its most striking feature is a logarithmic singularity, signifying physically a phase transition (ferromagnetic Curie point). Translated into communication theory, it can be said that at a certain critical strength of the intersymbol correlations, as measured by the parameter λ, there occurs a sudden collapse of transmission rate to a very low value, $dH/d\lambda$ becoming infinite at a single point.[14]

Conclusion

Much of what we have said has already been pointed out by others.[6,15] However, the basic mathematical identity of these two fields has had, thus far, very little influence on the development of either. There is an inevitable difference in detail, because the applications are so different; but we should at least develop a certain area of common language, so that a worker in one field can decide quickly whether work in the other has a bearing on his problems.

We suggest that one way of doing this is to recognize that the partition function, for many decades the standard avenue through which calculations in statistical mechanics are "channeled," is equally fundamental to communication theory. Even within communication theory, there are advantages to be had by adopting this terminology

[13] G. F. Newell and E. W. Montroll, "On the theory of the Ising model of ferromagnetism," *Rev. Mod. Phys.*, vol. 25, pp. 353–389; April, 1953.

[14] This type of message structure strongly resembles that occurring in certain styles of music, where strong correlations appear after an interval of 2^n bars, n being a small integer. This phenomenon of collapse in transmission rate then has some amusing implications, which we leave for the reader to develop.

[15] A referee kindly informs me that the following reference also contains material along the lines discussed here: Apostel, Mandelbrot, and Morf, "Linguistic statistique macroscopique" in "Logique, Langage et Theorie de L'information," Presses Universitaires de France, Paris, France, pp. 1–78; 1957.

and notation as standard. For example, expressions of the form $\sum D^{-n_i}$, which occur repeatedly in coding theory, are really partition functions. The "rather algebraic" nature of this theory derives in part from the fact that often only one value of D is considered. If we generalize by setting $D = 2^{\lambda}$, with λ a continuously variable parameter, we have a true partition function, which has analytical properties useful in deriving theorems; indeed, this is just what McMillan[4] has done. A partition function $Z(\lambda)$ is, of course, the same as a generating function of the variable $t = 2^{-\lambda}$. However, from a general standpoint the partition function is a more powerful analytical tool because it remains single-valued under conditions where the generating function would develop an infinite number of Riemann surfaces.

The way in which the partition function varies for different values of λ often tells one the effect of some departure from ideal conditions. Thus, in the problem treated above, we see from inspection of Fig. 2 that in the case of small alphabets it is essential to encode in such a way that the mean word length is held close to the optimal value; while in a large alphabet the mean word length can vary widely with very little effect on attainable transmission rate.

On the Use of Laguerre Polynomials in Treating the Envelope and Phase Components of Narrow-Band Gaussian Noise*

IRVING S. REED†

Summary—**The joint probability density of the envelope of a Gaussian process at two different times is expanded by the use of Hardy's identity into a series involving Laguerre polynomials. It is shown how this result may be used to estimate the cross-correlation function of the output of two quite general envelope-distorting filters. A generalization of this result, involving the use of the associated Laguerre polynomials, is obtained and applied to the calculation of a cross-correlation function which involves both the phase and envelope of the process at two points in time.**

IF $x(t)$ is a stationary narrow-band gaussian process with a power spectrum centered on the angular frequency ω, then it is well known[1] that $x(t)$ can be represented by

$$x(t) = R(t) \cos (\omega t + \theta(t))$$

where $R(t)$ is the envelope and $\theta(t)$ is the phase modulation. Price[2] has shown that the autocorrelation function of $x(t)$ is given by

$$\varphi_x(\tau) = Ex(t)x(t + \tau) = \sigma^2 \rho \cos (\omega\tau + \psi)$$

where E is the expected value operator, σ^2 is the rms noise power, $0 \leq \rho < 1$ is the normalized envelope, and ψ is the phase of $\varphi_x(\tau)$ as a function of τ.

The joint probability density of $R(t_1)$, $R(t_2)$, $\theta(t_1)$ and $\theta(t_2)$ is given by[1]

$$P(R_1, R_2, \theta_1, \theta_2) = \frac{R_1 R_2}{4\pi^2 \sigma^4 (1 - \rho^2)} \cdot \exp\left(-\frac{1}{2\sigma^2(1 - \rho^2)} \{R_1^2 + R_2^2 - 2\rho R_1 R_2 \cos(\theta_2 - \theta_1 - \psi)\}\right) \quad (1)$$

where $R_k = R(t_k)$, $\theta_k = \theta(t_k)$ for $(k = 1, 2)$. Rice[1] obtained the joint probability density of R_1 and R_2,

$$P(R_1, R_2) = \frac{R_1 R_2}{\sigma^4(1 - \rho^2)} I_0\left(\frac{\rho R_1 R_2}{\sigma^2(1 - \rho^2)}\right) \cdot \exp\left(-\frac{1}{2\sigma^2(1 - \rho^2)} \{R_1^2 + R_2^2\}\right), \quad (2)$$

from $P(R_1, R_2, \theta_1, \theta_2)$ by performing the integrations over θ_1 and θ_2 where I_0 is the modified Bessel function of the first kind. By Hardy's identity[3]

* Manuscript received by the PGIT, February 18, 1959. The work reported here was performed at the Lincoln Lab., a center for research operated by Mass. Inst. of Tech. with the joint support of the U. S. Army, Navy and Air Force.

† Staff member, Lincoln Lab., M.I.T., Lexington, Mass.

[1] S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 24, sec. 3.7, p. 46; July, 1944.

[2] R. Price, "A note on the envelope and phase-modulator components of narrow-band gaussian noise", IRE TRANS. ON INFORMATION THEORY, vol. IT-1, pp. 9–13; September, 1955.

[3] G. H. Hardy, "Summation of a series of polynomials of Laguerre," *J. London Math. Soc.*, vol. 7, pp. 138–139; 1932.

$$\frac{(yt)^{-1/2\alpha}}{1-t} I_\alpha\left\{\frac{2\sqrt{xyt}}{1-t}\right\} \exp\left(-\frac{t}{1-t}\{x+y\}\right)$$

$$= \sum_{n=0}^{\infty} L_n^{(\alpha)}(x) L_n^{(\alpha)}(y) t^n n!/\Gamma(\alpha+n+1) \qquad (3)$$

here

$$I_\alpha(u) = (\tfrac{1}{2}u)^\alpha \sum_{n=0}^{\infty} \frac{(\frac{1}{4}u^2)^n}{n!\Gamma(\alpha+n+1)}$$

the modified Bessel function of order α and

$$L_n^{(\alpha)}(u) = \frac{e^u u^{-\alpha}}{n!}\frac{d^n}{du^n}(e^{-u}u^{n+\alpha})$$

the generalized Laguerre polynomial of order α, we ow many express $P(R_1, R_2)$ in a series of terms involving he Laguerre functions of zero order, $L_n^{(0)}(u) \equiv L_n(u)$. etting

$$x = \frac{R_1^2}{2\sigma^2}, \qquad y = \frac{R_2^2}{\sigma^2} \quad \text{and} \quad \rho^2 = t$$

n (2), we have

$$P(R_1, R_2) = \frac{2\sqrt{xy}}{\sigma^2}\left[\frac{1}{1-t} I_0\left(\frac{2\sqrt{xyt}}{1-t}\right) \exp\left(-\frac{x+y}{1-t}\right)\right]$$

$$= \frac{2\sqrt{xy}}{\sigma^2}\left[\frac{1}{1-t} I_0\left(\frac{2\sqrt{xyt}}{1-t}\right)\right.$$

$$\left. \cdot \exp\left(-\frac{t}{1-t}\{x+y\}\right)\right] e^{-(x+y)}$$

$$= \frac{2\sqrt{xy}}{\sigma^2} \sum_{n=0}^{\infty} L_n(x) L_n(y) t^n e^{-(x+y)}$$

$$= \frac{R_1 R_2}{\sigma^2} \sum_{n=0}^{\infty} L_n\left(\frac{R_1^2}{2\sigma^2}\right) L_n\left(\frac{R_2^2}{2\sigma^2}\right) \rho^{2n} \exp\left\{-\frac{R_1^2+R_2^2}{2\sigma^2}\right\}. \qquad (4)$$

This expression is analogous to Cramer's expansion[4] of he two-dimensional joint gaussian distribution in terms f Hermite polynomials. Eq. (4) does not seem to have een observed before except by Levin[5] whose formula s in error. Levin's result has the associated Laguerre unction $L_n^{(1/2)}(x)$ in place of the proper Laguerre func- ion $L_n(x)$.

Suppose, now, that we have two envelope-distorting lters, one with output $G_1(R(t))$ and the other with output $_2(R(t))$. We are interested in the cross correlation of one f these outputs with the other delayed by a time τ. hat is, we wish to calculate

$$\Gamma_{12}(\tau) = EG_1(R(t_1))G_2(R(t_2)) = EG_1(R_1)G_2(R_2)$$

$$\equiv \int_0^\infty \int \frac{R_1 R_2}{\sigma^2(1-\rho^2)} I_0\left(\frac{\rho R_1 R_2}{\sigma^2\sqrt{1-\rho^2}}\right)$$

$$\cdot \exp\left\{-\frac{1}{2\sigma^2(1-\rho^2)}[R_1^2+R_2^2]\right\}$$

$$\cdot G_1(R_1)G_2(R_2)\, dR_1\, dR_2. \qquad (5)$$

[4] H. Cramer, "Mathematical Methods of Statistics," Princeton niversity Press, Princeton, N. J., p. 133; 1946.

[5] B. R. Levin, "The Theory of Noise Processes," Moscow Press, Ioscow, U. S. S. R., p. 310; 1957, (in Russian).

Proceeding formally, we have by (4)

$$\Gamma_{12}(\tau) = \frac{1}{\sigma^4} \sum_{n=0}^{\infty} \rho^{2n} \cdot\left[\int_0^\infty R_1 \exp\left\{-\frac{R_1^2}{2\sigma^2}\right\} L_n\left(\frac{R_1}{2\sigma^2}\right) G_1(R_1)\, dR_1\right] \cdot\left[\int_0^\infty R_2 \exp\left\{-\frac{R_2^2}{2\sigma^2}\right\} L_n\left(\frac{R_2}{2\sigma^2}\right) G_2(R_2)\, dR_2\right].$$

If we let

$$R_1 = \sigma\sqrt{2x} \quad \text{and} \quad R_2 = \sigma\sqrt{2y},$$

then

$$\Gamma_{12}(\tau) = \frac{1}{\sigma^2} \sum_{n=0}^{\infty} \rho^{2n} g_n^{(1)} g_n^{(2)} \qquad (6)$$

where

$$g_n^{(1)} = \int_0^\infty e^{-x} L_n(x) G_1(\sigma\sqrt{2x})\, dx,$$

and

$$g_n^{(2)} = \int_0^\infty e^{-y} L_n(y) G_2(\sigma\sqrt{2y})\, dy$$

are the Fourier coefficients (in the general mathematical sense) of the Fourier expansions of $G_1(\sigma\sqrt{2x})$ and $G_2(\sigma\sqrt{2y})$, respectively, in terms of Laguerre polynomials.

It is well known[6] that a function $f(x)$ where $0 \le x \le \infty$ may be expanded in a Fourier series of Laguerre poly. nomials of the form

$$\sum f_n L_n(x)$$

where

$$f_n = \int_0^\infty e^{-x} L_n(x) f(x)\, dx$$

if the integral

$$\int_0^\infty |\, f(x)\,|^2 e^{-x}\, dx < \infty$$

and exists in the Lebesgue sense. Thus if we assume that $G_1(\sigma\sqrt{2x})$ and $G_2(\sigma\sqrt{2y})$ are such that

$$\int_0^\infty |\, G_1(\sigma\sqrt{2x})\,|^2 e^{-x}\, dx = \int_0^\infty |\, G_1(R)\,|^2 \exp\left\{-\frac{R^2}{2\sigma^2}\right\} \frac{R\,dR}{\sigma^2} < \infty \qquad (7)$$

$$\int_0^\infty |\, G_1(\sigma\sqrt{2y})\,|^2 e^{-x}\, dx = \int_0^\infty |\, G_2(R)\,|^2 \exp\left\{-\frac{R^2}{2\sigma^2}\right\} \frac{R\,dR}{\sigma^2} < \infty$$

[6] G. Szego, "Orthogonal Polynomials," American Math. Soc. New York, N. Y., chs. III and V; 1939.

and exist in the Lebesgue sense, we are assured that both functions have Fourier expansions in Laguerre polynomials which converge to the original functions in the mean-square sense. Moreover, condition (6) may be used in conjunction with the identity[6]

$$\int_0^\infty e^{-x}L_n^2(x)\,dx = 1$$

and the Schwarz inequality for integrals to justify the change of the order of summation and integration we used to derive (6). This justification obtains[7] by showing

$$\sum_{n=0}^{\infty}\rho^{2n}\int_0^\infty\int |\, e^{-x-y}L_n(x)L_n(y)G_1(\sigma\sqrt{2x})G_2(\sigma\sqrt{2y})\,|\,dx\,dy$$

$$\le \sum_{n=0}^{\infty}\rho^{2n}\left(\int_0^\infty\int e^{-x-y}L_n^2(x)L_n^2(y)\,dx\,dy\right)^{1/2}$$

$$\cdot\left(\int_0^\infty\int e^{-x-y}\,|\,G_1\,\sigma\sqrt{2x})\,|^2\,|\,G_2\,\sigma\sqrt{2y})\,|^2\,dx\,dy\right)^{1/2}$$

$$= \frac{1}{1-\rho^2}\left(\int_0^\infty e^{-x}\,|\,G_1(\sigma\sqrt{2x})\,|^2\,dx\right)^{1/2}$$

$$\cdot\left(\int_0^\infty e^{-y}\,|\,G_2(\sigma\sqrt{2y})\,|^2\,dy\right)^{1/2} < \infty.$$

Evidently condition (7) is equivalent to the condition that the processes from the envelope-distorting filters have finite second moments, a condition which is generally true in practice.

Before we consider examples, let us calculate a more general expression which will include (5) as a special case. We will calculate

$$\Gamma_{12}^{(m)}(\tau) = EG_1(R_1)G_2(R_2)e^{im(\theta_2-\theta_1)} \tag{8}$$

where R_1, R_2, θ_1, θ_2 are defined as in (1) and G_1 and G_2 are defined as in (5). The quantity in (8) can be regarded as the cross correlation of the complex process $z_1(t) = G_1(R(t))\ e^{i\theta(t)}$ with conjugate of the process $z_2(t) = G_2(R(t))\ e^{i\theta(t)}$ but at the delay time τ. Evidently, by (1)

$$\Gamma_{12}^{(m)}(\tau)$$

$$= \int_0^\infty\int\int_0^{2\pi}\int G_1(R_1)G_2(R_2)e^{im(\theta_2-\theta_1)}P(R_1, R_2, \theta_1, \theta_2)$$

$$\cdot dR_1\,dR_2\,d\theta_1\,d\theta_2 = \frac{e^{im\psi}}{\sigma^4(1-\rho^2)}\int_0^\infty\int G_1(R_1)G_2(R_2)$$

$$\cdot R_1R_2I_m\left(\frac{\rho R_1R_2}{\sigma^4(1-\rho^2)}\right)\exp\left\{-\frac{R_1^2+R_2^2}{2\sigma^2(1-\rho^2)}\right\}dR_1\,dR_2$$

$$= e^{im\psi}\int_0^\infty\int G_1(\sigma\sqrt{2x})G_2(\sigma\sqrt{2y})$$

$$\cdot\left[I_m\left(\frac{2\sqrt{xyt}}{1-t}\right)\exp\left\{-\frac{(x+y)t}{1-t}\right\}\right]e^{-(x+y)}\,dx\,dy$$

which, by (3), formally becomes

$$= e^{im\psi}\int_0^\infty\int G_1(\sigma\sqrt{2x})G_2(\sigma\sqrt{2y})(xyt)^{m/2}$$

$$\cdot\sum_{n=0}^{\infty}L_n^{(m)}(x)L_n^{(m)}(y) \tag{9}$$

$$\frac{t^n n!}{\Gamma(m+n+1)}e^{-(x+y)}\,dx\,dy$$

$$= e^{im\psi}\rho^m\sum_{n=0}^{\infty}\rho^{2n}g_{n,m}^{(1)}g_{n,m}^{(2)}\frac{n!}{\Gamma(m+n+1)}$$

where $L_n^{(m)}(x)$ is the associated Laguerre polynomial

$$L_n^{(m)}(x) = \frac{e^x x^{-m}}{n!}\frac{d^n}{dx^n}(e^{-x}x^{n+m})$$

and $g_{n,m}^{(1)}$ and $g_{n,m}^{(2)}$ are the integral coefficients

$$g_{n,m}^{(1)} = \int_0^\infty e^{-x}x^{m/2}G_1(\sigma\sqrt{2x})L_n^{(m)}(x)\,dx$$

$$g_{n,m}^{(2)} = \int_0^\infty e^{-y}y^{m/2}G_2(\sigma\sqrt{2y})L_n^{(m)}(u)\,dy.$$

It is well known[6] that the coefficients $g_{n,m}^{(k)}$ for $k = 1, 2$ are the Fourier coefficients of the functions $G_k(\sigma\sqrt{2x})$ when expanded in a Fourier series of associated Laguerre polynomials of the form

$$G_k(\sigma\sqrt{2x}) \sim \sum_{n=0}^{\infty}g_{n,m}^{(k)}x^{m/2}L_n^{(m)}(x)\frac{n!}{\Gamma(m+n+1)}$$

where $k = 1, 2$. Clearly, the change of the order of operations needed to obtain (9) is justified if we again assume condition (7). Finally, formula (9) includes (6) as a special case since

$$g_{n,0}^{(k)} = g_n^{(k)}$$

for $(k = 1, 2)$.

Our principle example will be based on the following integral which will be proved in the Appendix,

$$\int_0^\infty e^{-t}t^{a-1}L_n^b(t)\,dt = \frac{\Gamma(a)}{n!}\frac{\Gamma(1-a+b+n)}{\Gamma(1-a+b)} \tag{10}$$

where $Re\ a > 0$ and where $L_n^b(t)$ is the generalized Laguerre polynomial

$$L_n^b(x) = \frac{e^x x^{-b}}{n!}\frac{d^n}{dx^n}(e^{-x}x^{n+b}).$$

If we let

$$G_1(R) = R^\alpha \quad \text{and} \quad G_2(R) = R^\beta$$

where α and β may be complex with real parts greater than -2, then by (9) and (10)

[7] E. C. Titchmarsh, "Theory of Functions," Oxford University Press, Cambridge, Eng., p. 45; 1932.

$$g_{n,m}^{(1)} = \int_0^\infty e^{-x} x^{m/2} (\sigma\sqrt{2x})^\alpha L_n^{(m)}(x)\, dx$$

$$= (\sigma\sqrt{2})^\alpha \frac{\Gamma\left(\frac{m+\alpha}{2}+1\right)}{n!} \frac{\Gamma\left(\frac{m-\alpha}{2}+n\right)}{\Gamma\left(\frac{m-\alpha}{2}\right)},$$

nd similarly

$$g_{n,m}^{(2)} = (\sigma\sqrt{2})^\beta \frac{\Gamma\left(\frac{m+\beta}{2}+1\right)}{n!} \frac{\Gamma\left(\frac{m-\beta}{2}+n\right)}{\Gamma\left(\frac{m-\beta}{2}\right)}$$

o that

$$\overline{R_1^\alpha R_2^\beta e^{im(\theta_2-\theta_1)}}$$

$$= e^{im\psi} \rho^m (\sigma\sqrt{2})^{\alpha+\beta} \frac{\Gamma\left(\frac{m+\alpha}{2}+1\right)\Gamma\left(\frac{m+\beta}{2}+1\right)}{m!}$$

$$\cdot \frac{\Gamma(m+1)}{\Gamma\left(\frac{m-\alpha}{2}\right)\Gamma\left(\frac{m-\beta}{2}\right)}$$

$$\cdot \sum_{n=0}^\infty \rho^{2n} \frac{\Gamma\left(\frac{m-\alpha}{2}+n\right)\Gamma\left(\frac{m-\beta}{2}+n\right)}{n!\Gamma(m+1+n)}$$

$$= e^{im\psi} \rho^m (\sigma\sqrt{2})^{\alpha+\beta}$$

$$\cdot \frac{\Gamma\left(\frac{m+\alpha}{2}+1\right)\Gamma\left(\frac{m+\beta}{2}+1\right)}{m!}$$

$$\cdot {}_2F_1\left(\frac{m-\alpha}{2}, \frac{m-\beta}{2}; m+1; \rho^2\right) \qquad (11)$$

where ${}_2F_1(a, b; c; z)$ is the hypergeometric function defined by

$${}_2F_1(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^\infty \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)} \frac{z^n}{n!}.$$

When $\alpha = \beta = \nu$ with ν real, (11) may be shown to be equivalent to a result first obtained by Middleton[8] for the correlation function of the m'th harmonic of the output of a ν'th law non-linear device into which narrow-band gaussian noise is fed. If $\alpha = \beta = 0$, (11) is the m-th Fourier coefficient of the probability density for $\theta_2 - \theta_1$, first obtained by MacDonald.[9] If $\alpha = r$, $\beta = s$ and $m = 0$ where r and s are positive integers, then (11) is the formula for the moments of joint envelope distribution (2). Finally, it is of some interest to consider a special case of (11) when α and β are complex. If $\alpha = -i\delta$ and $\beta = i\delta$ where δ is real and $m = 0$, then (11) reduces to

$$Ee^{i\delta \log R_2/R_1} = \left| \Gamma\left(1 + \frac{i\delta}{2}\right) \right|^2 {}_2F_1\left(\frac{i\delta}{2}, -\frac{i\delta}{2}; 1; \rho^2\right). \qquad (12)$$

This is a cross-correlation function of two functions of the envelope $R(t)$ at the respective times t_1 and t_2 which is independent of the noise power σ^2. Moreover, we have calculated, incidentally, the characteristic function of the random variable $\log R_2/R_1$. Since the right side of (12) is real and positive, we have

$$E\,(\log R_2/R_1)^{2n+1} = 0,$$

and

$$E\,(\log R_2/R_1)^{2n}$$

$$= (-1)^n \frac{d^{2n}}{d\delta^{2n}} \left\{ \left| \Gamma\left(1 + \frac{i\delta}{2}\right) \right|^2 {}_2F_1\left(\frac{i\delta}{2}, -\frac{i\delta}{2}; 1; \rho^2\right) \right\} \Bigg|_{\delta=0}$$

as the moments of the random variable $\log R_2/R_1$.

Appendix

The integral in (10) is a special case of the formula

$${}_2F_1(a, b; c; y) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-t} t^{a-1} {}_1F_1(b; c; yt)\, dt$$

given by Truesdall.[10] We set $y = 1$, and use the well-known identities[11]

$$L_n^b(y) = \binom{b+n}{b} {}_1F_1(-n; b+1; y)$$

and

$${}_2F_1(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

to obtain formula (10). Since $L_n^b(t)$ is a polynomial, it is easy to see that the integral exists when $Re\ a > 0$.

[8] D. Middleton, "Some general results in the theory of noise through nonlinear devices," *Quart. Appl. Math.*, vol. 5; 1948.

[9] P. K. C. MacDonald, "Some statistical properties of random noise," *Proc. Camb. Phil. Soc.*, vol. 45; 1949.

[10] C. Truesdall, "A Unified Theory of Special Functions," Princeton University Press, Princeton, N. J., p. 112; 1948.

[11] E. T. Copson, "Theory of Functions of a Complex Variable," Oxford University Press, Cambridge, Eng., Chapter X; 1935.

Interchannel Correlation in a Bank of Parallel Filters*

J. GALEJS† AND W. COWAN†

Summary—The first-order effects of interchannel noise correlation on the fasle alarm and incorrect dismissal probabilities are computed for a bank of parallel RLC filters by expanding the envelope distribution of the n filter outputs in a power series.

The correction to the false alarm probability due to noise correlation is found to decrease with increasing threshold-to-rms-noise ratios. If the filter-separation-to-filter-bandwidth ratios are larger than 0.2, it is less than 15 and 0.2 per cent for threshold-to-rms-noise ratios exceeding 12 and 14 db respectively.

The correction to the incorrect dismissal probability, which is computed by considering the signal output of three contiguous filters, increases with increasing threshold-to-rms-noise and signal-to-threshold ratios. Even for filter separations larger than the filter bandwidth, it may be in excess of 100 per cent if the threshold-to-rms-noise ratio exceeds 12 db and the signal-to-threshold ratio is larger than 1.2.

Introduction

MANY existing radar systems use banks of parallel filters prior to signal detection. Such filters act as coherent integrators or IF doppler filters. The determination of the false alarm or incorrect dismissal probabilities, which may be regarded as absolute criteria of system performance in a prescribed signal and noise environment, is complicated by the fact that noise outputs of the different filters are partially correlated. Because of the partial overlap of the filter pass bands, the noise outputs of adjacent filters will have a certain degree of similarity which affects the statistics of the filter outputs. In the simplest computational procedure, the partial noise correlation of the filter outputs is ignored and the envelope distribution of the outputs of n filters is computed as the product of the envelope distributions of the individual filters. If the partial noise correlation is taken into account, the envelope distribution of the n filter output is described by an n-fold integral, which cannot be evaluated in closed form. If the integrand is factorized after a power series expansion, the evaluation of this integral reduces to the evaluation of a product of n single integrals. The first-order effect of correlation between the filter outputs can be obtained by considering the first non-zero terms in this power series expansion.

Such a procedure has been applied for computing the false alarm rates of idealized coherent integrators, which integrate with uniform weighting the envelopes of the two phase-quadrature noise components.[1] However, no numerical results showing the magnitude of the interchannel correlation effects were given except for a statement that these effects are small.[2] A similar computational procedure can be used to approximate the amplitude distribution in the outputs of n parallel RLC filters if noise or noise plus signal are applied to the filter inputs. A series expansion in powers of correlation coefficients results in a factorized distribution that can be readily integrated to compute the false alarm and incorrect dismissal probabilities.

The Approximate Envelope Distribution of n Filter Outputs

The envelope distribution of the filter-outputs is computed using techniques already described.[1,3,4] Therefore, only a brief outline of the mathematical development will be included in this paper.

The system under consideration consists of n parallel RLC filters, each followed by an envelope detector and a threshold device as indicated in Fig. 1. The filters have a half-power bandwidth ω_d and are tuned to frequencies ω_i. The noise voltage output of a single filter which has been split into in-phase and phase-quadrature components can be represented as

$$V_i(t) = V_{ci}(t) \cos \omega_i t + V_{si}(t) \sin \omega_i t. \tag{1}$$

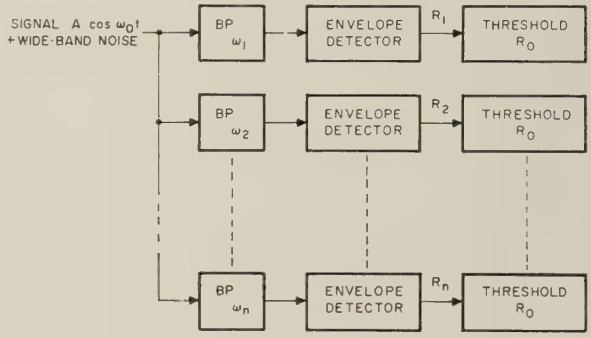


Fig. 1—Block diagram of the system.

Assuming that the noise components have zero mean Gaussian amplitude distributions, the averages of the output components of two RLC filters tuned to frequencies ω_i and ω_k are

$$\rho_{ik} = \langle V_{ci} V_{ck} \rangle = \langle V_{si} V_{sk} \rangle, \tag{2}$$

$$\mu_{ik} = \langle V_{ci} V_{sk} \rangle = -\langle V_{si} V_{ck} \rangle. \tag{3}$$

* Manuscript received by the PGIT, February 4, 1959.

† Appl. Res. Lab., Sylvania Electric Products Inc., Waltham, Mass.

[1] K. S. Miller and R. I. Bernstein, "An analysis of coherent integration and its application to signal detection," IRE TRANS. ON INFORMATION THEORY, vol. IT-3, pp. 237–248; December, 1957.

[2] *Op. cit.*, p. 245.

[3] S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, pp. 282–332; July, 1944; and *Bell Sys. Tech. J.*, vol. 24, pp. 46–156; January, 1945.

[4] D. Middleton, "Some general results in the theory of noise through nonlinear devices," *Quart. Appl. Math.*, vol. 5, pp. 445–498; January, 1948.

The averages ρ_{ik} and μ_{ik} represent the individual elements of the correlation matrix of a $2n$-dimensional Gaussian distribution that characterizes V_{cj} and V_{sj} of n filter outputs.[1,3,4]

In the absence of a signal, the envelope of a filter output can be computed as the absolute magnitude of $V_j(t)$ in (1). If a signal of amplitude A and of frequency ω_0 is applied to the filter bank, signal components of amplitude A_j will appear in the j'th filter. The amplitude A_j depends on the frequency separation

$$\omega_{dj} = \omega_0 - \omega_j \,. \tag{4}$$

The signal components A_j modify both the in-phase and phase-quadrature components of the filter output. The j'th filter output can also be described by an amplitude R_j and phase angle θ_j. The variables R_j and θ_j are thus functions of V_{sj}, V_{cj}, and of A_j. Alternately, the noise output components V_{cj} and V_{sj} that are characterized by a $2n$-dimensional Gaussian distribution can be related to A_j, R_j, θ_j. By substituting A_j, R_j, and θ_j in the appropriate functional relations for V_{cj} and V_{sj} in the $2n$-dimensional Gaussian distribution and using the required Jacobian, one obtains a probability distribution $W(R_1, \theta_1, \cdots, R_n, \theta_n)$ of the variables $R_1, \theta_1, \cdots, R_n, \theta_n$. The joint envelope distribution of the n filter outputs $R_1, \cdots, R_n$ is given by the n-fold integral

$$p(R_1, \cdots, R_n) = \int_0^{2\pi} \cdots \int_0^{2\pi} W(R_1, \theta_1, \cdots, R_n, \theta_n)\, d\theta_1 \cdots d\theta_n \,. \tag{5}$$

The complexity of the W function prohibits a direct evaluation of this integral. However, one can expand $p(R_1, \cdots, R_n)$ in powers of

$$g_{ik} = \sqrt{\mu_{ik}^2 + \rho_{ik}^2} \,, \tag{6}$$

which reduces products like $f_1(\theta_i)\, f_2(\theta_k)$ to sums $h_1(\theta_i) + h_2(\theta_k)$ in the exponents of the W function in (5). As a consequence, the n-fold integral (5) reduces to a product of n single integrals, each between the limits of 0 to 2π. Designating the envelope distribution of n outputs with all g_{ik}'s set equal to zero by $p(0)$, (5) can be expanded as[1]

$$p(R_1, \cdots, R_n) = p(0) + \sum_{\substack{i,k=1 \\ k>i}}^{n} \frac{\partial}{\partial g_{ik}} [p(0)] g_{ik} + \frac{1}{2} \sum_{\substack{i,k=1 \\ k>i}}^{n} \sum_{\substack{r,s=1 \\ s>r}}^{n} \frac{\partial^2}{\partial g_{ik} \partial g_{rs}} [p(0)] g_{ik} g_{rs} + \cdots \,. \tag{7}$$

The derivatives of $p(0)$ are computed by setting all g_{ik}'s except the ones to be differentiated equal to zero before the differentiation. If the series expansion is terminated with the quadratic terms, the variables that are related by the non-zero g_{ik}'s are characterized by at most a 6 by 6 correlation matrix.[5] For a single signal appearing in the center of the i'th channel, (7) reduces to

$$p(R_1, R_2, \cdots R_n) = \prod_{i=1}^{n} p_i(R_i) + \sum_{\substack{i,k=1 \\ k>1}}^{n} D_{ik} g_{ik}^2 \left[\prod_{\substack{j=1 \\ j \neq i,k}}^{n} p_j(R_j) \right] + \sum_{k=i+1}^{m} E_{ikr} g_{ik} g_{ri} \left[\prod_{\substack{j=1 \\ j \neq i,k,r}}^{n} p_j(R_j) \right] \tag{8}$$

where

$$p_i(R_i) = \frac{R_i}{\psi} \exp\left[-\frac{R_i^2 + A_i^2}{2\psi} \right] I_0\left(\frac{R_i A_i}{\psi}\right), \tag{9}$$

$$\begin{aligned} 2D_{ik} &= \left[\frac{\partial^2 p(R_i, R_k)}{\partial g_{ik}^2} \right]_{g_{ik}=0} \\ &= \frac{2R_iR_k}{\psi^4} \exp\left[-\frac{1}{2\psi} (R_i^2 + R_k^2 + A_i^2 + A_k^2) \right] \cdot \left\{ \frac{1}{4\psi^2} (R_i^2 R_k^2 \right. \\ &\quad + R_i^2 A_k^2 + A_i^2 A_k^2 + R_k^2 A_i^2) I_0\left(\frac{R_i A_i}{\psi}\right) I_0\left(\frac{R_k A_k}{\psi}\right) \\ &\quad - \frac{1}{2\psi^2} (R_i R_k^2 A_i + R_i A_i A_k^2) I_0\left(\frac{R_k A_k}{\psi}\right) I_1\left(\frac{R_i A_i}{\psi}\right) \\ &\quad - \frac{1}{2\psi^2} (R_i^2 R_k A_k + R_k A_i^2 A_k) I_0\left(\frac{R_i A_i}{\psi}\right) I_1\left(\frac{R_k A_k}{\psi}\right) \\ &\quad + \frac{1}{\psi^2} R_i R_k A_i A_k I_1\left(\frac{R_i A_i}{\psi}\right) I_1\left(\frac{R_k A_k}{\psi}\right) \\ &\quad - \frac{1}{2\psi} (R_i^2 + R_k^2 + A_i^2 + A_k^2) I_0\left(\frac{R_i A_i}{\psi}\right) I_0\left(\frac{R_k A_k}{\psi}\right) \\ &\quad + \left(\frac{R_i A_i}{\psi}\right) I_1\left(\frac{R_i A_i}{\psi}\right) I_0\left(\frac{R_k A_k}{\psi}\right) \\ &\quad + \left(\frac{R_k A_k}{\psi}\right) I_0\left(\frac{R_i A_i}{\psi}\right) I_1\left(\frac{R_k A_k}{\psi}\right) \\ &\quad \left. + I_0\left(\frac{R_i A_i}{\psi}\right) I_0\left(\frac{R_k A_k}{\psi}\right) \right\}, \end{aligned} \tag{10}$$

$$\begin{aligned} 2E_{irk} &= \left[\frac{\partial^2 p(R_r, R_i, R_k)}{\partial g_{ri} \partial g_{ik}} \right]_{g_{ri}=g_{ik}=g_{rk}=0} = \frac{R_r R_i R_k}{4\psi^7} \\ &\quad \cdot \exp\left[-\frac{1}{2\psi} (R_r^2 + R_i^2 + R_k^2 + A_r^2 + A_i^2 + A_k^2) \right] \\ &\quad \cdot \left[R_r I_1\left(\frac{R_r A_r}{\psi}\right) - A_r I_0\left(\frac{R_r A_r}{\psi}\right) \right] \\ &\quad \cdot \left[R_k I_1\left(\frac{R_k A_k}{\psi}\right) - A_k I_0\left(\frac{R_k A_k}{\psi}\right) \right] \\ &\quad \cdot \left[0.5 R_i^2 I_2\left(\frac{R_i A_i}{\psi}\right) - R_i A_i I_1\left(\frac{R_i A_i}{\psi}\right) \right. \\ &\quad \left. + 0.5 A_i^2 I_0\left(\frac{R_i A_i}{\psi}\right) \right], \end{aligned} \tag{11}$$

and where I_n is the modified Bessel function of the first

[5] If i,k,r,s have 3 distinct values, there is a 6×6 correlation matrix. If i,k,r,s have 4 distinct values, there are two 4×4 matrices.

kind and order n. The symbols r and m are given by

$$r = 2i - k \tag{12}$$

and

$$m = \begin{cases} n & \text{if } i \geq 0.5(n+1), \\ 2i - 1 & \text{if } i \leq 0.5(n+1). \end{cases} \tag{13}$$

The variable r depends on i and k. E of (11) is therefore characterized by only two independent parameters, i and k.

False Alarm Probability

The probability of a false alarm α in the output of a bank of n filters is equal to the probability that at least one filter output exceeds a prescribed threshold R_0 in the absence of signal. Thus

$$\alpha = 1 - \int_0^{R_0} \cdots \int_0^{R_0} p(R_1, R_2, \cdots R_n)\, dR_1 \cdots dR_n, \tag{14}$$

where the probability distribution of the filter output envelopes $p(R_1, R_2, \cdots R_n)$ is computed with all the signal amplitudes A_i set equal to zero. $A_i = 0$ results in zero E_{ik}. D_{ik} reduces to[1]

$$D_{ik} = \frac{R_i R_k}{\psi^4}\left[1 - \frac{R_i^2}{2\psi}\right]\left[1 - \frac{R_k^2}{2\psi}\right] \cdot \exp\left[-\frac{1}{2\psi}(R_i^2 + R_k^2)\right]. \tag{15}$$

The false alarm probabilities obtained by considering or by ignoring the interchannel correlation effects are denoted by α_{corr} or by $\alpha_{\text{no corr}}$ respectively. Evaluation of (14) shows that

$$\frac{\alpha_{\text{no corr}} - \alpha_{\text{corr}}}{\alpha_{\text{no corr}}} = \frac{0.25x^2 \exp(-x)[1 - \exp(-0.5x)]^{n-2}}{1 - [1 - \exp(-0.5x)]^n} \cdot \sum_{i=1}^{n-1} \frac{n-i}{1 + i^2 y^2}, \tag{16}$$

where

$$x = \frac{R_0^2}{\psi} \tag{17}$$

and

$$y = \frac{\Delta\omega}{\omega_d} = \frac{(\omega_{i+1} - \omega_i)}{\omega_d}. \tag{18}$$

The changes in false alarm probability due to noise correlation are computed numerically for different threshold-squared-to-mean-square-noise ratios x, filter-separation-to-filter-bandwidth ratios y, and numbers of filters n. The upper limit of the summation over i was varied for n = constant in the computations leading to the curves of Fig. 2 in order to show the effect of noise correlation between increasingly more distant filters. The upper limit of the summation was equal to $(n - 1)$ in the plots of Figs. 3–7; these plots take noise correlation between all n filters into account.

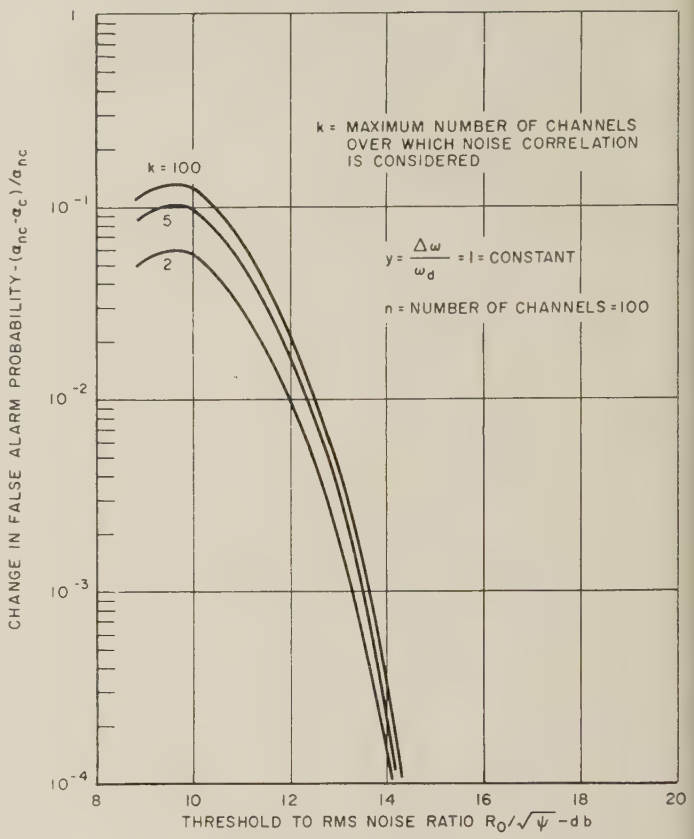


Fig. 2—Change in false alarm rate with the number of channels over which noise correlation is considered.

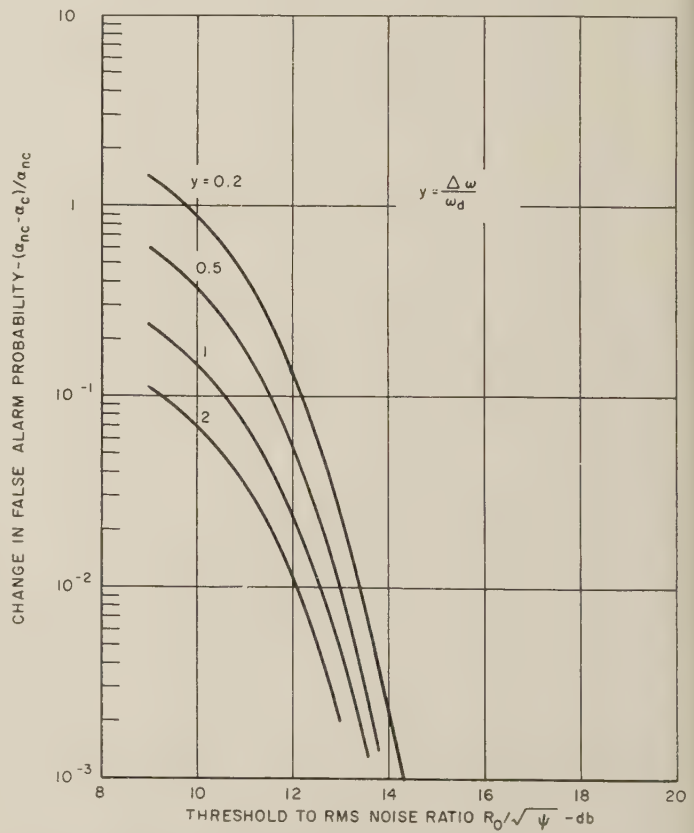


Fig. 3—Change in false alarm rates for different amounts of channel overlap, $n = 20$.

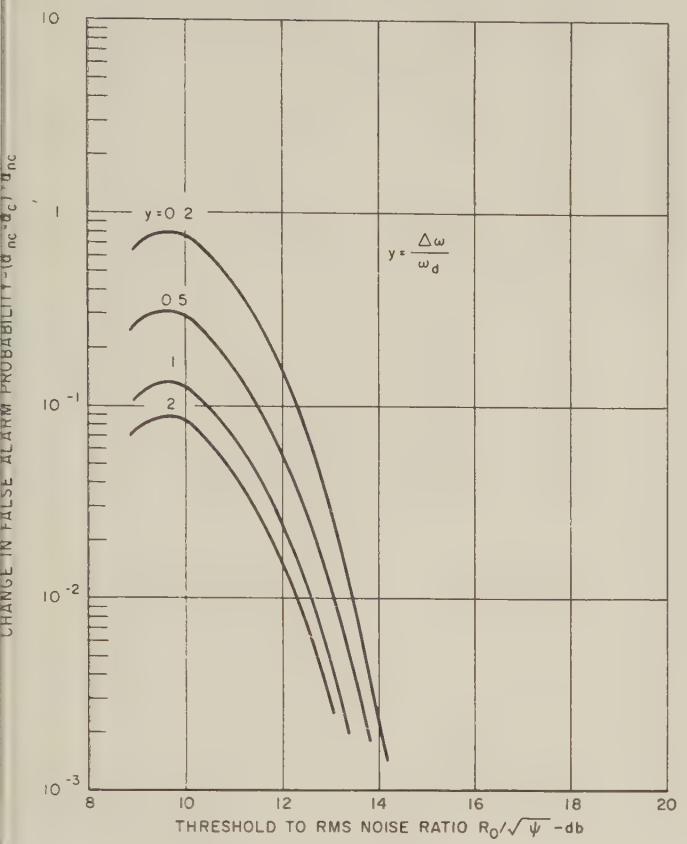


Fig. 4—Change in false alarm rates for different amounts of channel overlap, $n = 100$.

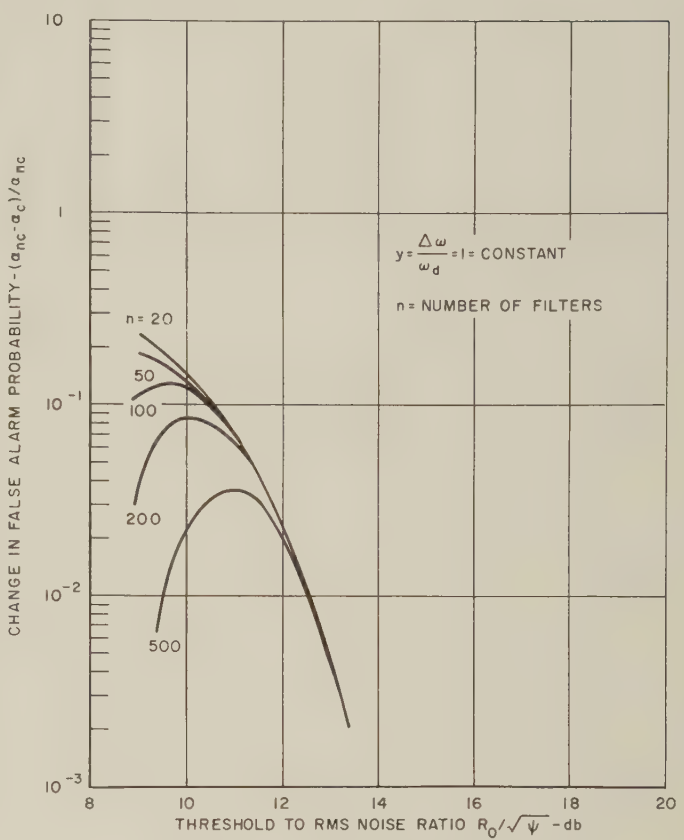


Fig. 6—Change in false alarm rate with different number of filters of constant bandwidth and constant filter separation.

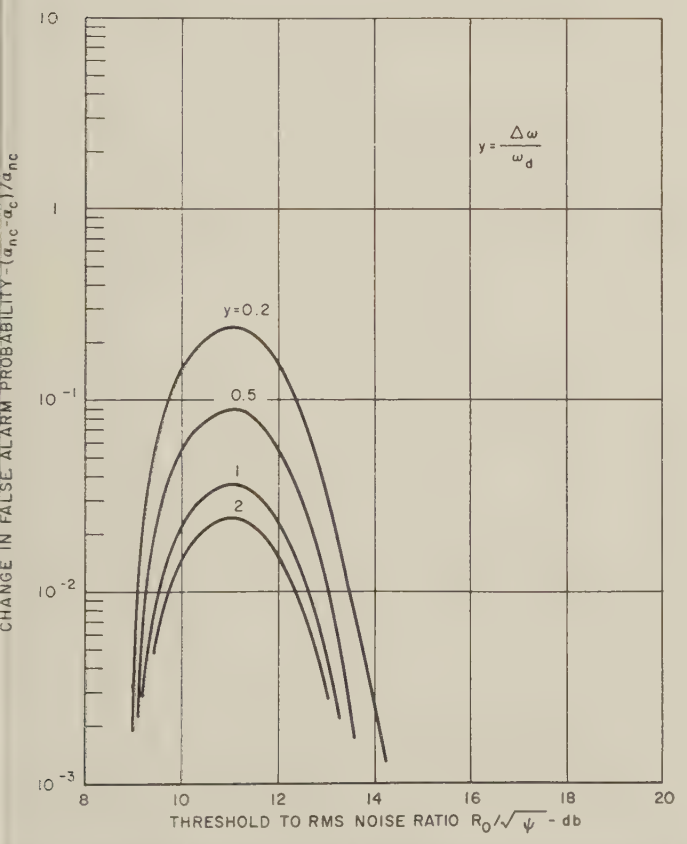


Fig. 5—Change in false alarm rates for different amounts of channel overlap, $n = 500$.

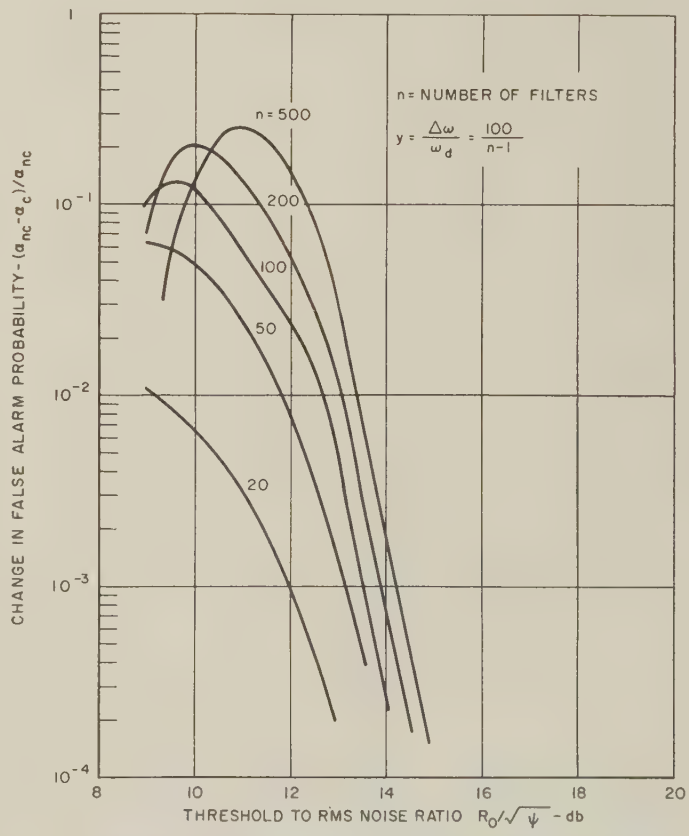


Fig.7—Change in false alarm rate with different number of filters of constant bandwidth covering the same frequency band.

Probability of Incorrect Dismissal

The probability of incorrect dismissal at a prescribed threshold R_0 is equal to the probability that all of the filter outputs remain below a threshold R_0, if a signal of amplitude A is applied to the filter bank. Thus

$$\beta = \int_0^{R_0} \cdots \int_0^{R_0} p(R_1, \cdots R_n)\, dR_1 \cdots dR_n. \tag{19}$$

The value of β obtained by substituting $p(R_1, \cdots R_n)$ of (8) into (19) is designated as β_{corr}. Substituting the joint probability distribution with neglected correlation effects into (19), one obtains the incorrect dismissal probability $\beta_{\text{no corr}}$. The ratio of these two incorrect dismissal probabilities is clearly

$$\frac{\beta_{\text{corr}}}{\beta_{\text{no corr}}} = 1 + \sum_{s=2}^{n} \sum_{r=1}^{s-1} \frac{g_{rs}^2 \int_0^{R_0} \int_0^{R_0} D_{rs}\, dR_r\, dR_s}{K_s(R_0) K_r(R_0)} + \sum_{k=i+1}^{m} g_{ik}^2 \frac{\int_0^{R_0} \int_0^{R_0} \int_0^{R_0} E_{ikr}\, dR_i\, dR_r\, dR_k}{K_i(R_0)[K_k(R_0)]^2}, \tag{20}$$

where g_{rs}, D_{rs}, E_{ikr}, r, and m are given by (6), (10), (11), (12), and (13) respectively, and where

$$K_i(R_0) = \int_0^{R_0} p_i(R_i)\, dR_i. \tag{21}$$

The above integrals and the summations have been evaluated numerically by means of a digital computer. The plots in Figs. 8–17 show the change in the incorrect dismissal probability for different threshold-squared-mean-square-noise ratios x, filter-separation-to-filter-bandwidth ratios y, signal-to-threshold ratios ($z = A/R_0$), and number of filters n.

Discussion

False Alarm Probabilities

The first-order correction to the false alarm probability α is computed by evaluating (16) for different threshold-to-rms-noise ratios ($R_0/\sqrt{\psi} = \sqrt{x}$), filter-separation-to-bandwidth ratios ($\Delta\omega/\omega_d = y$), and numbers of filters n.

The upper limit of the summation in (16) determines the number of adjacent channels k, over which noise correlation is considered. With the limit equal to $(n - 1)$, noise correlation over all n filters is taken into account. If the summation is approximated by the $i = 1$ term, only the effects of noise correlation between adjacent filters are considered. It is shown in Fig. 2 that for $\Delta\omega = \omega_d$ noise correlation over a few filter channels ($k = 5$) gives almost the same correction to the false alarm probability α as noise correlation over all possible filter pairs ($k = n$).

Figs. 3 to 5 show the correction to the false alarm probability α with increasing threshold-to-rms-noise ratios $R_0/\sqrt{\psi}$ at different filter separations y. The number of filters n is constant in each figure. Increased filter separations y decrease the noise correlation between adjacent channels, which decreases the α correction. The

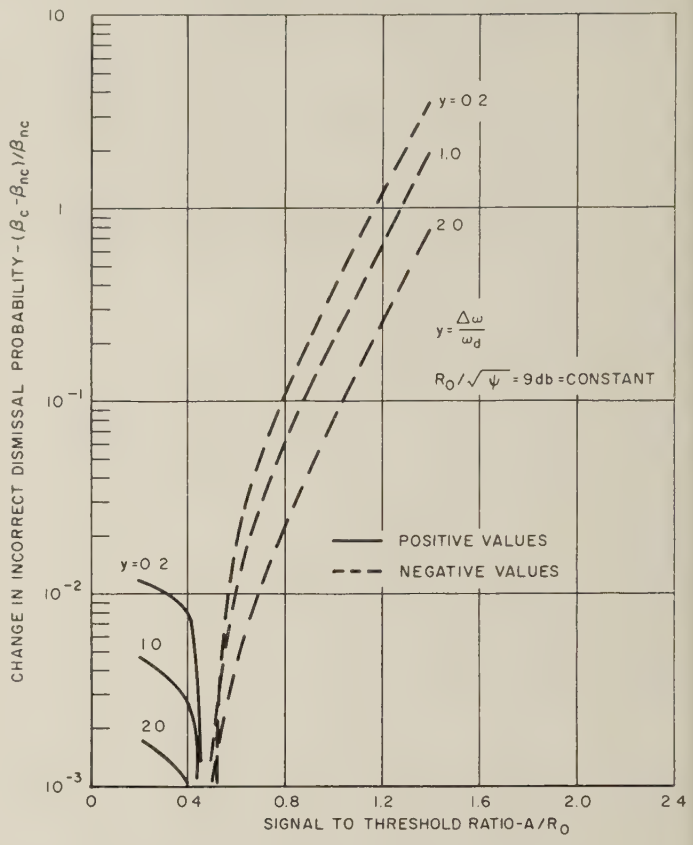


Fig. 8—Change in incorrect dismissal probability for different signal-to-threshold ratios, 3 filters, signal output only in the center one.

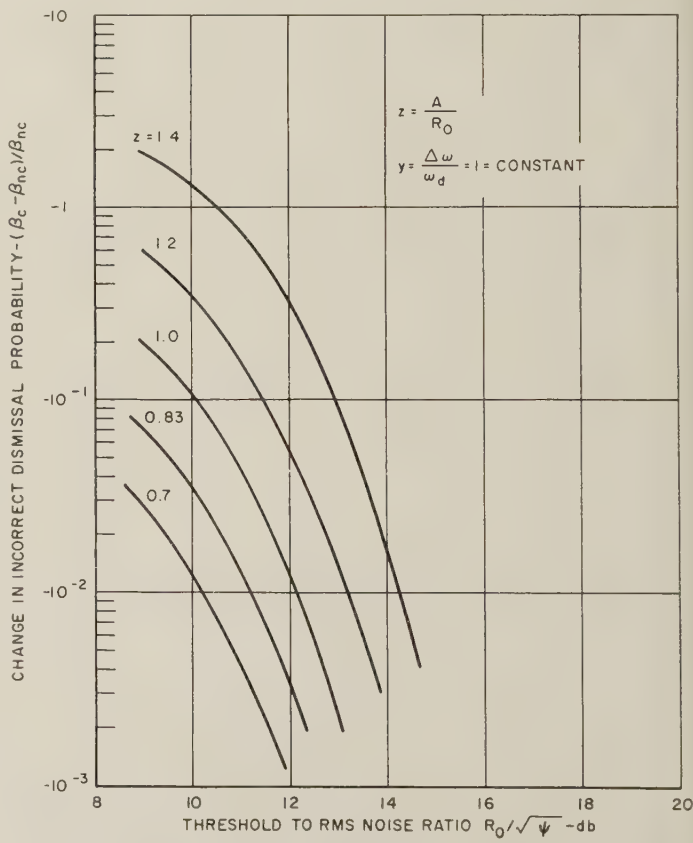


Fig. 9—Change in incorrect dismissal probability for different Threshold-to-rms-noise ratios, 3 filters, signal output only in the center one, filter separation $y = 1$.

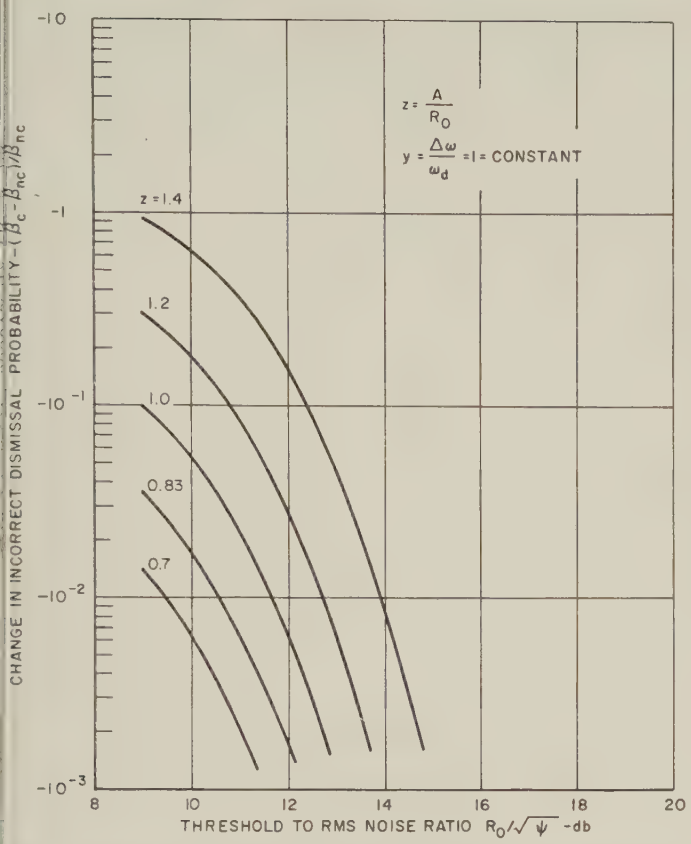


Fig. 10 Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, 5 filters, signal output only in the center one, filter separation $y = 1$.

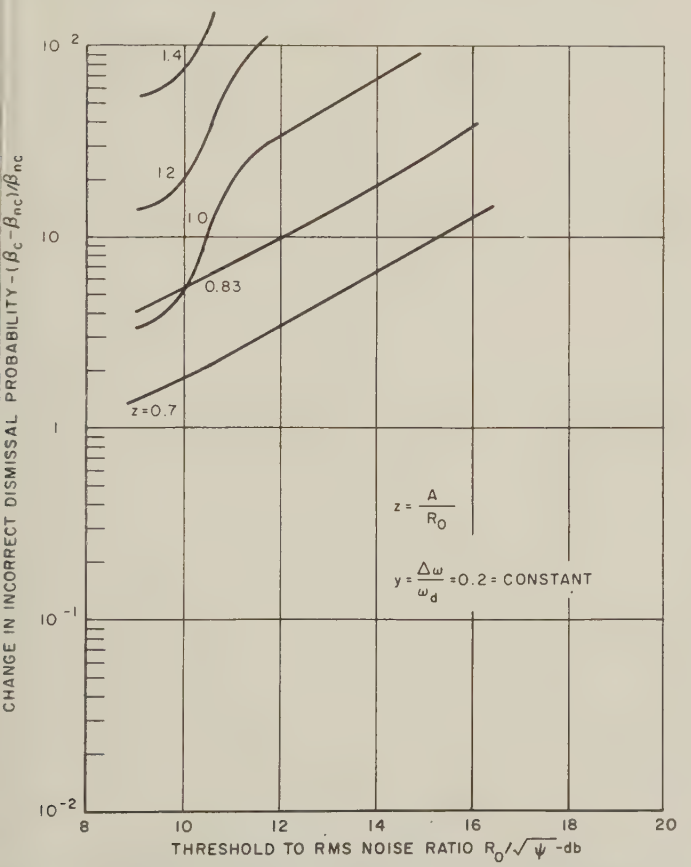


Fig. 11—Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, signal output in all 3 filters, filter separation $y = 0.2$.

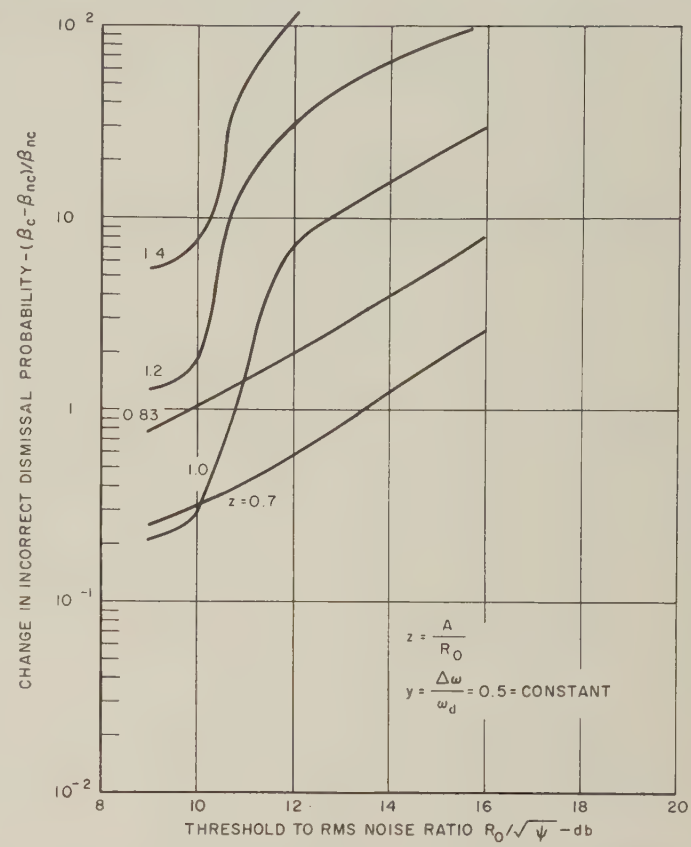


Fig. 12—Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, signal output in all 3 filters, filter separation $y = 0.5$.

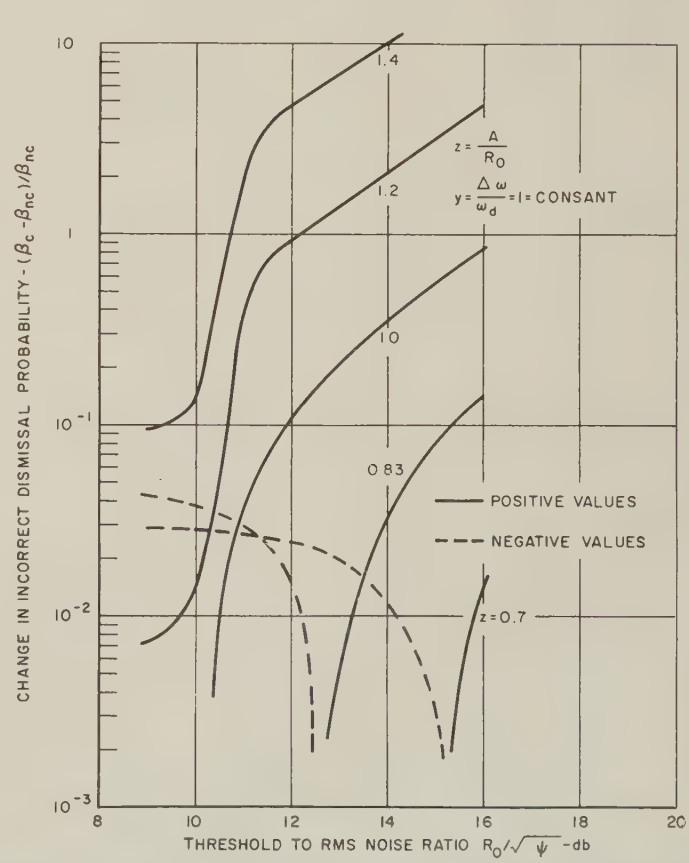


Fig. 13—Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, signal output in all 3 filters, filter separation $y = 1$.

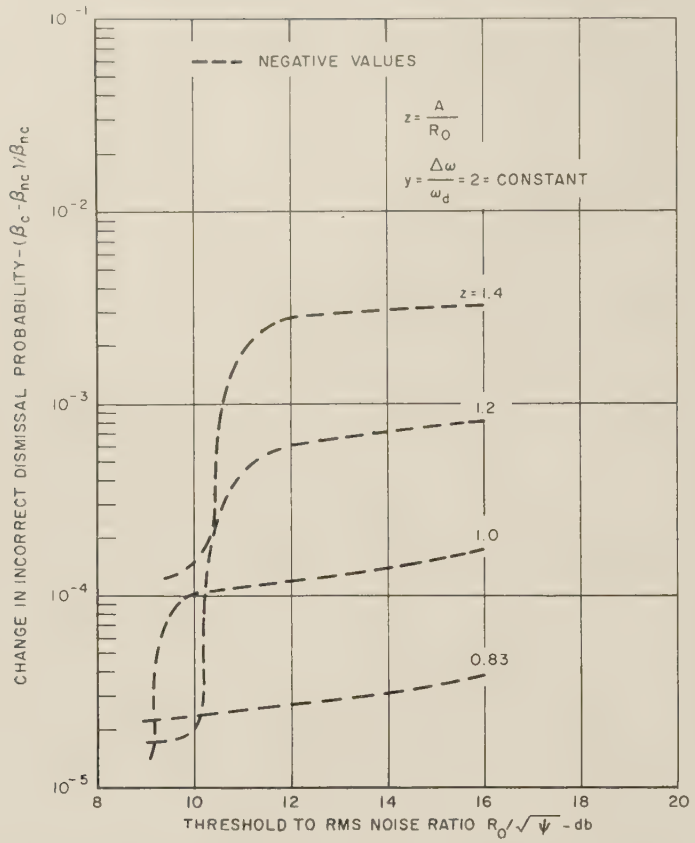


Fig. 14—Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, signal output in all 3 filters, filter separation $y = 2$.

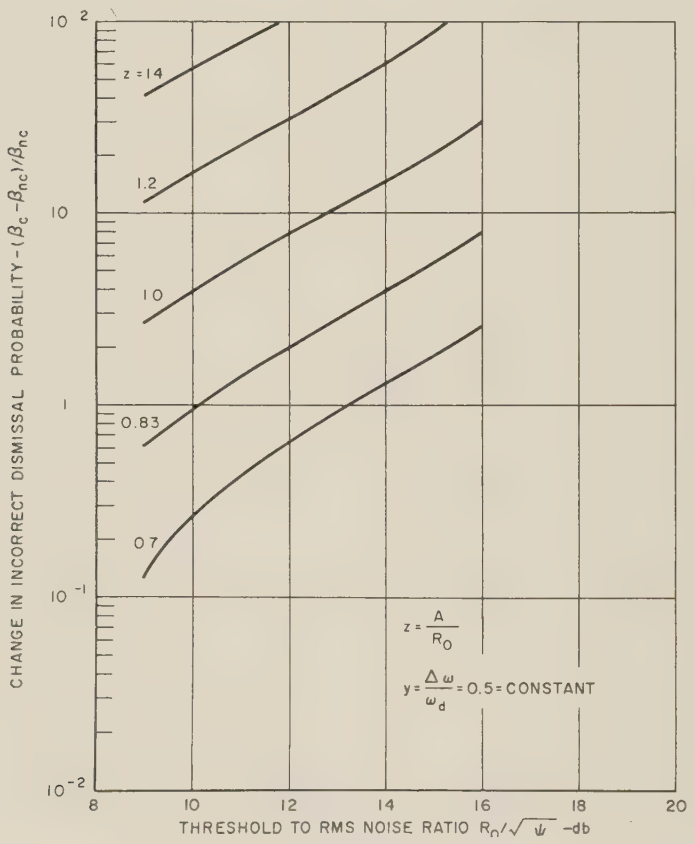


Fig. 15—Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, 5 filters, signal output in the 3 center ones, filter separation $y = 0.5$.

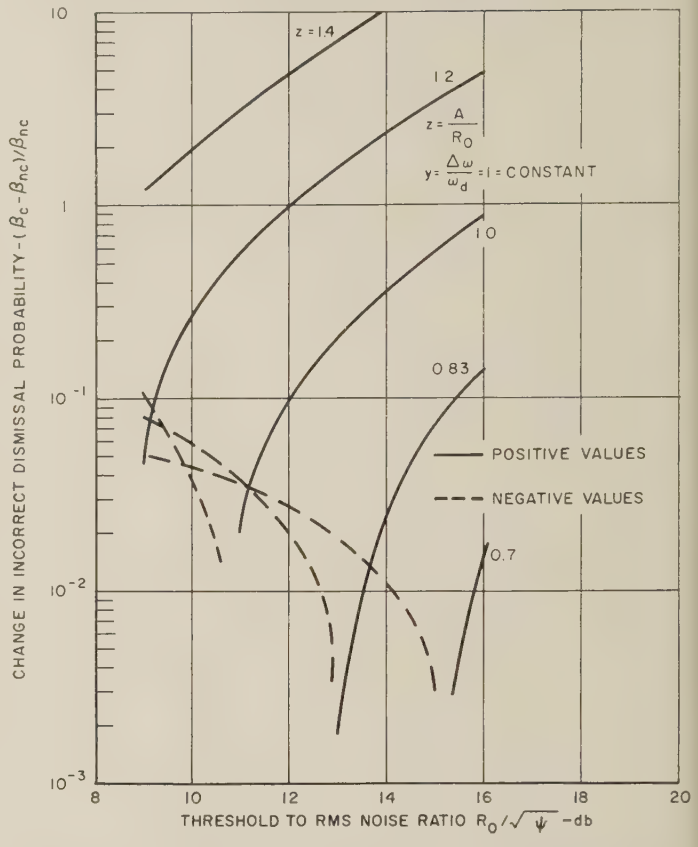


Fig. 16—Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, 5 filters, signal output in the 3 center ones, filter separation $y = 1$.

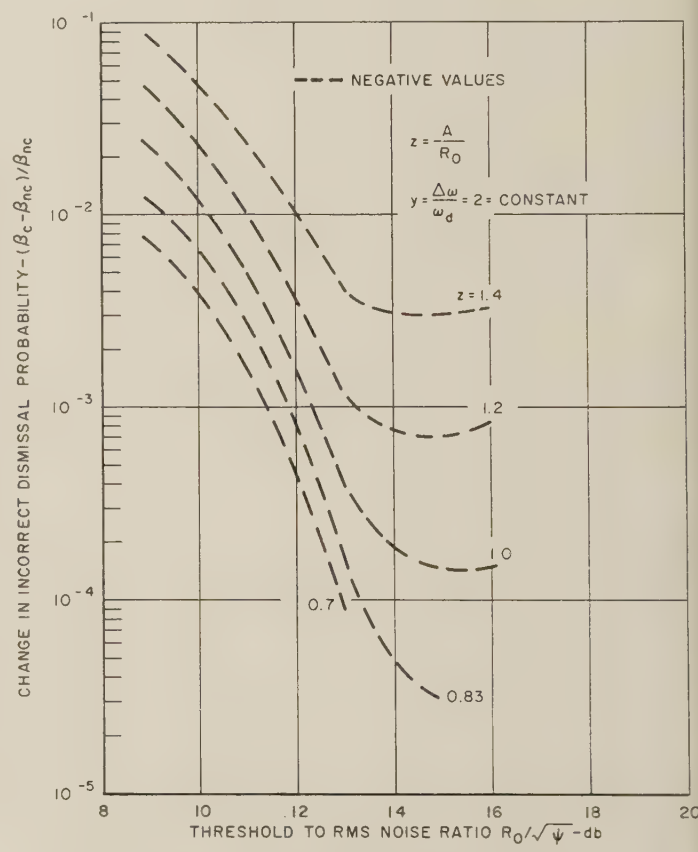


Fig. 17—Change in incorrect dismissal probability for different threshold-to-rms-noise ratios, 5 filters, signal output in the 3 center ones, filter separation $y = 2$.

correction to the false alarm probability α is small for large threshold-to-rms-noise ratios. It is less than 0.2 per cent for $R_0/\sqrt{\psi} = 14$ db, but it can be as high as 15 per cent for $R_0/\sqrt{\psi} = 12$ db. Increasing the number of filters n decreases the maximum value of the α correction while shifting it to slightly higher values of $R_0/\sqrt{\psi}$. This maximum may exceed unity for $n = 20$, but it is decreased to less than 25 per cent for $n = 500$. The maximum occurs for $n = 50$ below $R_0/\sqrt{\psi} = 9$ db, and for $n = 500$ at $R_0/\sqrt{\psi} = 11$ db.

Similar relations between the correction to the false alarm probability α and the number of filters n can be seen from Fig. 6, which is obtained by replotting the $y = 1$ curves of Figs. 3–5. Increasing n represents here an increased over-all filter bandwidth if the bandwidth ω_d and separation $\Delta\omega$ of the individual filters is kept constant.

The n dependency of the curves shown in Figs. 3–6 can be explained qualitatively by considering the false alarm probability α for small threshold-to-rms-noise ratios $R_0/\sqrt{\psi}$. The probability α for a single channel is nearly unity for sufficiently small $R_0/\sqrt{\psi}$ ratios. For $n \neq 1$, α is increased but remains less than unity regardless of noise correlation in the channels. This accounts for the small $(\alpha_{\text{no corr}} - \alpha_{\text{corr}})/\alpha_{\text{no corr}}$ values at small $R_0/\sqrt{\psi}$ ratios. With increasing $R_0/\sqrt{\psi}$, α is decreased but it remains closer to unity if n is large. Therefore, the possible α correction is decreased for n large.

When a constant frequency band is covered with an increasing number of filters n, the filter separation $\Delta\omega$ is decreased and the filter overlap increased if the filter bandwidth ω_d remains constant. The maxima of the α correction curves in Fig. 7 still shift to higher $R_0/\sqrt{\psi}$ ratios with increasing n but the magnitude of the α correction is increased with increasing n. This is an obvious consequence of the filter overlap which causes an increased noise correlation in adjacent filter channels.

Incorrect Dismissal Probabilities

The first-order correction to the incorrect dismissal probability β is computed by evaluating (20) for different threshold-to-rms-noise ratios $R_0/\sqrt{\psi} = \sqrt{x}$, filter-separation-to-filter-bandwidth ratios $\Delta\omega/\omega_d = y$, signal-to-threshold ratios $A/R_0 = z$, and numbers of filters n.

The first set of computations assumes the signal to be in the center channel of three or five parallel filters. The signal output of all other filters is neglected. Such an idealization may be prompted by an attempt to simplify the rather complex numerical calculations.

Fig. 8 shows that the correction to the incorrect dismissal probability,

$$\frac{\beta_{\text{corr}} - \beta_{\text{no corr}}}{\beta_{\text{no corr}}},$$

is positive for small signal amplitudes A. For very small signal amplitudes A,

$$\beta \approx 1 - \alpha \tag{22}$$

as seen from (16) and (21). This gives

$$\frac{\beta_{\text{corr}} - \beta_{\text{no corr}}}{\beta_{\text{no corr}}} \approx \frac{\alpha_{\text{no corr}} - \alpha_{\text{corr}}}{1 - \alpha_{\text{no corr}}} \tag{23}$$

which is positive according to the α curves discussed earlier. The β correction in Fig. 8 reverses sign and becomes negative for larger A/R_0 ratios. The total β correction will be (-1) for A/R_0 exceeding 2 and fully correlated noise in the filter channels. This is evident from the vector diagram in Fig. 18. For full noise correlation the noise vectors in the signal-plus-noise and in the noise-only channels are identical. Whenever the signal-plus-noise tends to be within the circle of radius R_0, the noise will be outside the circle. The probability β_{corr}, which is the probability that both signal-plus-noise and noise-only give an output of less than R_0, is equal to zero. Thus, the correction $(\beta_{\text{corr}} - \beta_{\text{no corr}})/\beta_{\text{no corr}}$ is equal to (-1) whenever $A > 2R_0$.

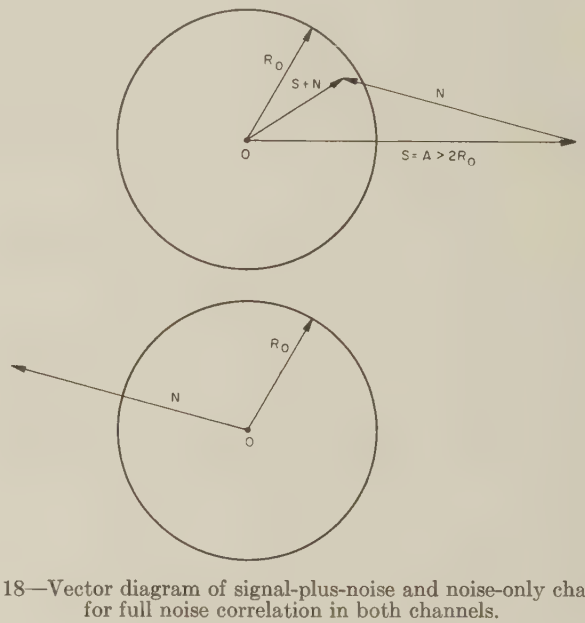


Fig. 18—Vector diagram of signal-plus-noise and noise-only channel for full noise correlation in both channels.

The correction to the incorrect dismissal probability β is shown in Figs. 9 and 10 for varying threshold-to-rms-noise ratios $R_0/\sqrt{\psi}$ and for different z and n values. Increasing z and n increases the amount of the required negative β correction.

The β corrections are recomputed for $n = 3$ and the signal in the center channel without neglecting the signal output of the two side channels. The curves plotted in Fig. 11 to 14 indicate that the β corrections increase with $R_0/\sqrt{\psi}$ contrary to the observations made in earlier figures. Also, first order corrections in excess of 100 times are noted for the larger A/R_0 and $R_0/\sqrt{\psi}$ ratios.

This can be readily explained by considering the limit of fully correlated noise in several filter channels. Fully correlated noise is obtained in completely overlapping filter channels. The signals in the output of such channels are identical. The incorrect dismissal probability β of any

number of such channels is equal to the probability β_0 of a single channel. In n channels containing uncorrelated noise and identical signals, the incorrect dismissal probability is β_0^n. The β correction becomes herewith

$$\frac{\beta_{\text{corr}} - \beta_{\text{no corr}}}{\beta_{\text{no corr}}} = \frac{1 - \beta_0^{n-1}}{\beta_0^{n-1}}. \tag{24}$$

The probability β_0 gets small for $(A - R_0)/\sqrt{\psi}$ large, which is the case for large A/R_0 and $R_0/\sqrt{\psi}$ ratios; the β correction reduces to

$$\frac{\beta_{\text{corr}} - \beta_{\text{no corr}}}{\beta_{\text{no corr}}} \approx \frac{1}{\beta_0^{n-1}}, \tag{25}$$

which increases with decreasing β_0 or with increasing A/R_0 and $R_0/\sqrt{\psi}$ ratios.

The curves plotted in Figs. 11–14 decrease with increasing filter separations y, but β corrections in excess of 100 per cent are possible even for $\Delta\omega > \omega_d$ with $R_0/\sqrt{\psi} >$ 12 db and $A/R_0 > 1.2$.

Five parallel filters with the signal located in the center channel are assumed in the computations leading to the plots of Figs. 15–17. Signal output of only the three center channels is considered; the signal output of the two outer channels is neglected. The $n = 5$ curves in Figs. 15–17 exhibit approximately the same large $R_0/\sqrt{\psi}$ values as the $n = 3$ curves in Figs. 11–14. For smaller $R_0/\sqrt{\psi}$ values the $n = 5$ curves exhibit either larger positive or larger negative values, with a sign change occurring at approximately the same $R_0/\sqrt{\psi}$ values as for the $n = 3$ curves.

Comparison of Figs. 9 and 10 with Figs. 11–17 indicates that neglecting the signal output of the filters that are off the signal frequency results in β correction factors of erroneous $R_0/\sqrt{\psi}$ dependency and even of wrong signs. Although the reference to Figs. 9 and 10 may be of little value in connection with practical decision problems, they are still included in this paper for illustration purposes.

Limitations of the Present Work

The curves plotted in Figs. 2–17 represent the first-order corrections to the false alarm and incorrect dismissal probabilities due to noise correlation in the different filter channels. This first-order correction can be expected to approximate the total correction only if the power series is rapidly convergent as in the case of small values for the first-order correction. The first-order terms may give a non-accurate indication of the total correction whenever the first-order correction is close to or even above unity. The calculation of higher-order corrections has not been attempted in this paper.

Application of Modular Sequential Circuits to Single Error-Correcting *P*-Nary Codes*

T. E. STERN† AND B. FRIEDLAND†

I. Introduction

IT is the purpose of this paper to present systematic methods for the construction of close-packed single error-correcting multi-level codes, coders and decoders. The emphasis in this paper will be laid on simple and efficient means of constructing the coders and decoders using linear modular sequential circuits.[1,2]

Error correcting and detecting codes for binary channels have been developed by Hamming,[3] Slepian,[4] Huffman,[5,6] and others. The work of Huffman is of special importance here because of his use of linear sequential filters[6] for coding and decoding. The modular sequential circuits to be described in this work are generalizations of Huffman's filters. Multi-level codes have been discussed by Golay[7]

* Manuscript received by the PGIT, March 11, 1959. This work was supported in part by the Department of the Navy, under Contract No. NONR 266(60).

† Department of Electrical Engineering, Columbia University, New York, N. Y.

[1] B. Friedland, "Linear modular sequential circuits," IRE TRANS. ON CIRCUIT THEORY, vol. CT-6, pp. 61–68; March, 1959.

[2] B. Friedland and T. E. Stern, "On Periodicity of States in Linear Modular Sequential Circuits," this issue, pp. 136–137.

[3] R. W. Hamming, "Error detecting and error correcting codes," *Bell Sys. Tech. J.*, vol. 29, pp. 147–160; April, 1950.

[4] D. Slepian, "A class of binary signaling alphabets," *Bell Sys. Tech. J.*, vol. 35, pp. 203–234; January, 1956.

[5] D. A. Huffman, "A linear circuit viewpoint on error-correcting codes," IRE TRANS. ON INFORMATION THEORY, vol. IT-2, pp. 20–28; September, 1956.

[6] D. A. Huffman, "The synthesis of linear sequential coding networks," "Information Theory," C. Cherry, ed., Academic Press, Inc., New York, N. Y., pp. 77–95; 1956.

[7] M. J. E. Golay, "Notes on digital coding," PROC. IRE, vol. 37, p. 637; June, 1949.

and Ulrich.[8] However, little attention has been given to the problem of finding simple means of realizing coders and decoders in the multi-level case.

II. Close-Packed Single Error-Correcting P-Nary Codes

Consider a block of l symbols, each symbol having p possible levels (say, the integers, 0, 1, $\cdots$, $p-1$). Assume that only single errors can occur in this block, and that each error may occur in w possible ways ($w \leq p-1$). Further, assume that each block contains k check digits, the remainder being arbitrary message digits. In a received and decoded message block, the condition "all check digits zero," is to correspond to a transmission without error; other combinations of values of the check digits are to correspond *uniquely* to the wl ways in which a single error can occur. Since there are $p^k - 1$ non-zero combinations of check digits, we have

$$wl \leq p^k - 1 \tag{1}$$

or

$$l \leq \frac{p^k - 1}{w}. \tag{2}$$

If condition (2) is fulfilled with the equals sign, we have a one-one correspondence between each possible single error and each possible non-zero combination of check digits. A close-packed systematic code will be defined as one in which condition (2) is fulfilled with the equals sign.

In the special case in which any symbol may be transformed to any other symbol by an error in transmission, we have for a close-packed code,

$$l = \frac{p^k - 1}{p - 1} \quad \begin{pmatrix} = \text{an integer for all integral} \\ \text{values of } p \text{ and } k \end{pmatrix}. \tag{3}$$

Clearly, for a given number of levels p there is a restricted set of block lengths which result in close-packed codes. (See Table I.) Of course, in a practical case, a block length only slightly shorter than the close-packed length would still result in a fairly efficient code. For reasons stated in Friedland,[1] only codes containing a prime number of levels are considered in this paper. (Binary codes, of course, are a special case.)

TABLE 1
Values of $l = p^k - 1/p - 1$

k \ p	2	3	5	7	11
1	1	1	1	1	1
2	3	4	6	8	12
3	7	13	31	57	133
4	15	40	156	400	1464

[8] W. Ulrich, "Non-binary error correction codes," *Bell Sys. Tech. J.*, vol. 36, pp. 1341–1388; November, 1957.

III. Linear Modular Sequential Circuits

Since the presentation which follows utilizes the device known as a linear modular sequential circuit (LMSC), a brief summary of the properties of this device is included here. For a more complete discussion of LMSC's see Friedland and Stern.[1,2]

A linear modular sequential circuit, modulo p, a prime, is a system comprising unit delays, amplifiers whose gains are integers $< p$, and summers, mod p. Such circuits can be treated by more or less direct extensions of most of the tools of linear circuit theory.

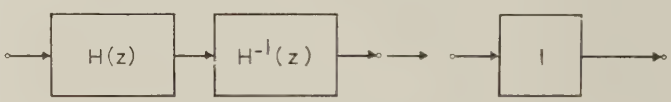
Any LMSC, with one input $x(n)$ and one output $y(n)$, can be characterized by a difference equation

$$y(n+k) + \alpha_1 y(n+k-1) + \cdots + \alpha_k y(n) = \beta_0 x(n+k) + \beta_1 x(n+k-1) + \cdots + \beta_k x(n) \tag{4}$$

over the modular field $GF(p)$. The modular field $GF(p)$ consists of the integers 0, 1, $\cdots$, $p-1$ together with the operations $\cdot$ and $+$, modulo p. If the difference equation is of order k, the LMSC is said to be of order k. Following Friedland,[1] we may, by the use of the Z-transform (or similarly, by the use of the delay operator), derive a transfer function for the system of (4) in the form

$$H(z) = \frac{\beta_0 z^k + \beta_1 z^{k-1} + \cdots + \beta_k}{z^k + \alpha_1 z^{k-1} + \cdots + \alpha_k}. \tag{5}$$

The response of this system to a unit input ("unit impulse response")[9] can be calculated from (5) in a number of ways, *e.g.*, long division. For a given $H(z)$, we may also construct an inverse circuit, $H^{-1}(z)$, such that the transfer function of the cascade combination is unity (See Fig. 1). Two circuits are mutually inverse if their transfer functions are reciprocals. (An inverse of a given circuit is not always realizable, however.)



$$H(z) = \frac{P(z)}{Q(z)}$$

P and Q are polynomials in z

$$H^{-1}(z) = \frac{Q(z)}{P(z)}$$

Fig. 1—Inverse filters.

From (4) or (5), we may construct a canonical flow graph representation of the circuit as shown in Fig. 2. Note that the feedback gains correspond to the coefficients of the denominator of the transfer function. Expressions

[9] It is convenient, but by no means necessary, to interpret the input and output time functions as (equally spaced) trains of impulses having magnitudes ("areas") equal to 1, 2, $\cdots$ $p-1$.

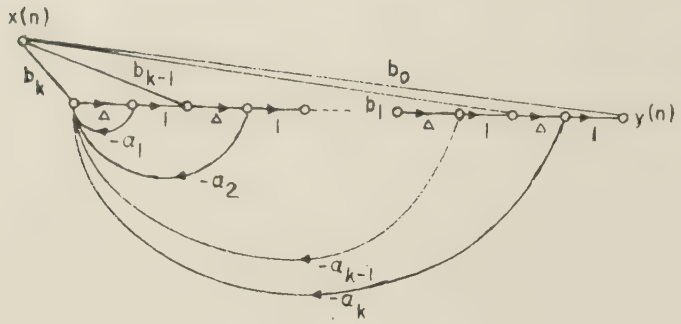


Fig. 2—Flow graph representation of LMSC.

for the b_i are

$$\begin{aligned} b_0 &= \beta_0 \\ b_1 + \alpha_1 b_0 &= \beta_1 \\ b_2 + \alpha_1 b_1 + \alpha_2 b_0 &= \beta_2 \\ &\vdots \\ b_k + \sum_{i=1}^{k} \alpha_i b_{k-i} &= \beta_k. \end{aligned} \tag{6}$$

IV. The Coding and Decoding System

Following the method of Huffman,[5] we may construct a coding and decoding system using two LMSC's as shown in Fig. 3. The system consists of a coding filter, $H(z)$; the noisy channel, the effect of which is represented by a noise sequence N introduced at a summing point; and a decoding filter $H^{-1}(z)$, which is the inverse of the coding filter. A decision mechanism operates on the output of $H^{-1}(z)$ to restore the distorted received message to its original form. A feature of this system is that the decision operation is extremely simple, as we shall subsequently demonstrate.

Operation of the system is as follows [it will be assumed in this section that $w = p - 1$, so that $l = (p^k - 1)/(p - 1)$]:

A message block of length l consisting of $(l - k)$ arbitrary message digits followed by k check digits (all zeros) constitutes the input sequence X to the coding filter. The output of the coding filter is, therefore,

$$U = HX.$$

[In the sequel, it is understood that capital letters, *e.g.*, $H(z)$, H, will refer to Z-transforms of sequences, and lower case letters, *e.g.*, $h(n)$, h, to the sequences themselves.]

The distorted message which is received by the decoder may be represented as $U' = HX + N$, where N is a sequence of length l whose i'th digit equals the magnitude of the error which has occurred in the i'th place of U. For example, the noise sequence N for a message whose fourth digit changed from 2 to zero, mod 5, would be $0\,0\,0\,3\,0\,0\,0\,0\cdots$

At the receiver, the distorted message U' is first operated on by the decoding filter to produce an output

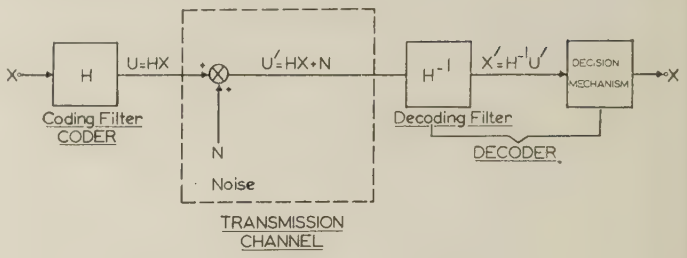


Fig. 3—Coding and decoding channel.

$$X' = H^{-1}U' = H^{-1}(HX + N) = X + H^{-1}N. \tag{7}$$

Eq. (7) indicates that X', the output of the decoding filter, consists of a digit-by-digit sum of the original message X and the response of the decoding filter to the noise impulse. (Only single noise impulses are considered.) Clearly, then, for no errors the message will be reproduced correctly at the output of the decoding filter. Otherwise, it will be necessary to subtract $H^{-1}N$ from X' to obtain the correct message. It is the function of the decision mechanism to determine the form of $H^{-1}N$, and to subtract this from X'.

The determination of $H^{-1}N$ is accomplished in the following manner. First, we observe that since the last k digits of X were originally zero, then from (7) the last k digits of X' must consist only of $H^{-1}N$ without any contribution from X. Assume that N corresponds to a single error of magnitude q occurring at time n. Then, by linearity, the sequence $H^{-1}N$ is simply q times the response of the filter H^{-1} to a unit impulse occurring at time n. Therefore, the last k digits of x' will always consist of a subsequence of k digits of the unit impulse response of H^{-1} or a constant multiple thereof. (Of course, in the case of error impulse occurring in one of the last k places, say the $(l - r)$th place, where $r < k$, the last k digits of x' will consist of $r + 1$ digits of the impulse response or its multiple, preceded by $k - r - 1$ zeros.)

Now, let us assume that by observing the subsequence contained in the last k digits of x' we can *uniquely* extrapolate backwards in time to determine the complete form of $H^{-1}N$. If this is possible for every subsequence which might occur, then we have solved the decoding problem. It is not difficult to deduce that such a unique extrapolation will always be possible if $h^{-1}(n)$ has the following two properties:

A) The sequence consisting of the first l digits of $h^{-1}(n)$ preceded by $k - 1$ zeros must comprise l subsequences of length k, each of which must be distinct.

B) None of these subsequences may be a constant multiple (mod p) of any other. (8)

If these two conditions hold, then the set of all subsequences and their multiples, as defined above, will correspond in a 1 — 1 fashion to the $p^k - 1$ ways in which a single error can occur, and therefore, the required unique extrapolation will be possible. The fundamental problem in the construction of the coding and decoding system is

the synthesis of a decoding filter having properties (8). This problem and its solution will be discussed in section VI. First, let us consider the following:

Example A

Let $p = 5$, $k = 2$, then $l = 6$.

From the theory of sec. VI, it is found that an appropriate set of filters is given by

$$H^{-1}(z) = \frac{z^2}{z^2 + 4z + 2}$$

$$H(z) = \frac{z^2 + 4z + 2}{z^2}.$$

It can be verified that

$$h^{-1}(n) = 1\ 1\ 4\ 2\ 4\ 0,$$
$$2\ 2\ 3\ 4\ 3\ 0,$$
$$4\ 4\ 1\ 3\ 1\ 0,$$
$$3\ 3\ 2\ 1\ 2\ 0;$$
$$1\ 1\ 4 \cdots$$

Observe that h^{-1} is periodic (the semi-colon denotes the end of the period of length 24), and that the period can be subdivided into four subsequences of length l, each of which is a constant multiple, mod 5, of the first. Inspection of the first six digits of h^{-1} (preceded by a zero) indicates that properties (8) are satisfied.

Consider a typical message sequence

$$x = 1\ 2\ 3\ 1\ 0\ 0.$$

Then

$$u = 1\ 1\ 3\ 2\ 0\ 2.$$

Let the noise sequence be

$$n = 0\ 3\ 0\ 0\ 0\ 0.$$

Then

$$u' = 1\ 4\ 3\ 2\ 0\ 2$$

and

$$x' = 1\ 0\ 1\ 3\ 1\ 2.$$

(These relations can be verified by standard transform techniques.)

Remark

A property of this system that can be observed from this example is that the information digits as well as the check digits of x become altered in passing through the coder.

To identify $H^{-1}N$ and extrapolate backwards, we examine the last two digits of x'. We note that this pair, 1 2, appears in the 4th and 5th places of the last subsequence of h^{-1}. But the last subsequence is simply three times the first, corresponding to the first six digits of the response of H^{-1} to an impulse of magnitude three. Thus, we may deduce that an error of three has occurred in the second position of x. Hence

$$H^{-1}N = 0\ 3\ 3\ 2\ 1\ 2.$$

Subtracting this from X', we obtain

$$\begin{array}{r} 1\ 0\ 1\ 3\ 1\ 2 \\ -\ 0\ 3\ 3\ 2\ 1\ 2 \qquad \text{mod } 5. \\ \hline x = 1\ 2\ 3\ 1\ 0\ 0 \end{array}$$

Before proceeding, let us examine the properties of the filter chosen for example A

$$\left.\begin{array}{l} 1)\ h^{-1}(n) \text{ is periodic with period } p^k - 1. \\ 2)\ H(z) \text{ exists and is realizable.} \\ 3)\ H^{-1}(z) \text{ and } H(z) \text{ are ratios of polynomials in } z \\ \text{of degree not exceeding } k. \end{array}\right\} \quad (9)$$

It will be shown in section VI that properties (9) imply properties (8), and imply, in addition, that h^{-1} consists of $p - 1$ disjoint and distinct subsequences of length $(p^k - 1)/(p - 1)$, each one of which is a multiple of the first.

We now summarize the steps in the construction of close-packed single error-correcting codes, coders and decoders (assume p a prime, and $w = p - 1$):

1) Select a block length l and a number of check digits k such that $l = (p^k - 1)/(p - 1)$.

2) Select a pair of coding and decoding filters possessing properties (9).

3) Construct each message $x(n)$ in the form of a block consisting of $l - k$ information digits followed by k zeros.

4) Obtain the coded message $u(n)$ as the response of the coding filter H to $x(n)$.

5) Obtain the decoded noisy message $x'(n)$ as the response of the decoding filter H^{-1} to $u'(n)$. If the last k digits are zeros, the message is correct. If not, proceed to step 6.

6) Write out the impulse response $h^{-1}(n)$ with commas separating the subsequences. Shift $x'(n)$ over $h^{-1}(n)$, until the last k digits of $x'(n)$ match a corresponding set in $h^{-1}(n)$. This will always be possible since h^{-1} will contain every non-zero combination of digits taken k at a time.

7) Of the set of digits appearing directly below x', subtract all those which follow the comma. (One and only one comma can appear preceding a digit directly under x'.) The resultant sequence is $x(n)$, the restored message.

An additional example is helpful in illustrating these seven steps.

Example B

1) Choose $p = 5$, $k = 3$, $l = 31$.

2) For this case we find that an appropriate filter is given by

$$H^{-1}(z) = \frac{z^3}{z^3 + z^2 + z + 3}.$$

It can then be verified by division that

$$h^{-1}(n) = 1\ 4\ 0\ 3\ 0\ 2\ 4\ 4\ 1\ 3\ 4\ 0\ 2\ 1\ 2\ 1\ 4\ 4\ 4\ 0\ 4\ 4\ 2\ 2\ 4\ 3\ 2\ 3\ 1\ 0\ 0, 2\ 3\ 0\ 1 \cdots$$

and that the sequence has the required properties.

3) Consider a typical message block,

$$x(n) = 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ 1\ 0\ 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ 1\ 0\ 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ 0\ 0\ 0\ .$$

4) The message, $u(n)$, at the output of the coding filter,

$$H(z) = (z^3 + z^2 + z + 3)/z^3, \qquad \text{is}$$

$$u(n) = 1\ 3\ 1\ 2\ 3\ 2\ 4\ 4\ 3\ 2\ 3\ 1\ 1\ 2\ 3\ 2\ 4\ 4\ 3\ 2\ 3\ 1\ 1\ 2\ 3\ 2\ 4\ 4\ 2\ 1\ 1\ 0\ 0 \cdots .$$

Let us assume that an error of 2 occurs in the 9th place, producing a zero in the 9th place of $u'(n)$.

5) The response of H^{-1} to $u'(n)$ is

$$x'(n) = 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ \underline{3\ 3\ 1\ 3\ 3\ 3\ 3\ 2\ 0\ 3\ 4\ 0\ 0\ 4\ 2\ 1\ 3\ 2\ 1\ 2\ 3\ 3\ 4} \cdots .$$

Note that all digits of $x'(n)$ starting with the ninth (as underlined) differ from those of $x(n)$.

6) Writing $h^{-1}(n)$ in a position to match the last k digits (and incidentally, all those that follow),

$$x\,(n) = 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ 3\ 3\ 1\ 3\ 3\ 3\ 3\ 2\ 0\ 3\ 4\ 0\ 0\ 4\ 2\ 1\ 3\ 2\ 1\ 2\ 3\ 3\ 4 \cdots$$

$$h^{-1}(n) = \cdots\ 3\ 2\ 3\ 1\ 0\ 0, 2\ 3\ 0\ 1\ 0\ 4\ 3\ 3\ 2\ 1\ 3\ 0\ 4\ 2\ 4\ 2\ 3\ 3\ 3\ 0\ 3\ 3\ 4 \cdots$$

$$x(n) = 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ 1\ 0\ 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ 1\ 0\ 1\ 2\ 3\ 4\ 0\ 4\ 3\ 2\ 0\ 0\ 0 \cdots .$$

7) The last sequence is the restored message.

V. The State Vector-Matrix Formulation

In the synthesis of the coding and decoding filters, a characterization of the LMSC in terms of a state vector and a characteristic matrix is more convenient than the transfer function characterization utilized in section II. The formulation which follows is completely equivalent to that used in the preceding discussion.

Referring to Fig. 2, we may represent the state of a k'th order circuit at time n as the column vector

$$\mathbf{s}(n) = \begin{bmatrix} s_1(n) \\ s_2(n) \\ \vdots \\ s_k(n) \end{bmatrix}.$$

We define

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

as the "null state," and it will be assumed that $s(0) = \mathbf{0}$, unless otherwise stated. Using the state concept, the operation of the circuit is described as follows

$$\mathbf{y}(n) = C\mathbf{s}(n) + D\mathbf{x}(n)$$

$$\mathbf{s}(n+1) = A\mathbf{s}(n) + B\mathbf{x}(n) \qquad (10)$$

where

$$[A] = \begin{bmatrix} -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{k-1} & -\alpha_k \\ 1 & 0\ 0 & \cdots & & \\ 0 & 1\ 0 & \cdots & & \\ \vdots & & & & \vdots \\ 0 & 0\ 0 & \cdots & 1 & 0 \end{bmatrix}. \qquad (11)$$

(This form of matrix is known as a "companion" form.)

$$|A - \lambda I| = \lambda^k + \alpha_1\lambda^{k-1} + \alpha_2\lambda^{k-2} + \cdots + \alpha_{k-1}\lambda + \alpha_k.$$

In the case of one input and one output,

$$\mathbf{x}(n) = x(n)$$

$$\mathbf{y}(n) = y(n)$$

$$[B] = \begin{bmatrix} b_k \\ b_{k-1} \\ \vdots \\ b_1 \end{bmatrix} \quad [C] = [00 \cdots 01] \quad [D] = b_0. \qquad (12)$$

The relations between the β_i's appearing in $H(z)$ and the b_i's appearing in Fig. 2 and in the matrices have been defined previously [see (6)].

It will be observed that the same polynomial appears throughout this discussion, first in the difference equation, then as the denominator of the transfer function, and now as the characteristic equation of A. The matrix A is defined as the "characteristic matrix" of the LMSC, and it will be seen that its characteristic equation plays a fundamental role in the coding problem.

Of particular interest here is the response of this system to a unit impulse. Since it is more convenient to deal with the unexcited system, we can, as in ordinary linear system theory, interpret the impulse response as the *homogeneous* response to some "virtual" initial state, $\mathbf{s}'(0)$. (We say "virtual" since the system has not actually started from this initial state. We have simply used $\mathbf{s}'(0)$ as an artifice to replace the effect of the initial impulse.) Using this interpretation, the equations for an unexcited system of order k become:

$$\begin{aligned} \mathbf{s}(n) &= A^n\mathbf{s}'(0) \\ y(n) &= s_k(n) \end{aligned} \qquad \mathbf{s}'(0) = \begin{bmatrix} b_{k-1} \\ b_{k-2} \\ \vdots \\ b_0 \end{bmatrix}. \tag{13}$$

The form of $\mathbf{s}'(0)$ can easily be verified by reference to Fig. 2.

The LMSC without excitation can be visualized simply as a shift register. The content of its left-hand place at time $n + 1$ is calculated as a linear combination (mod p) of the contents of all positions of the register at time n. Fig. 4 shows such a register in its virtual initial state.

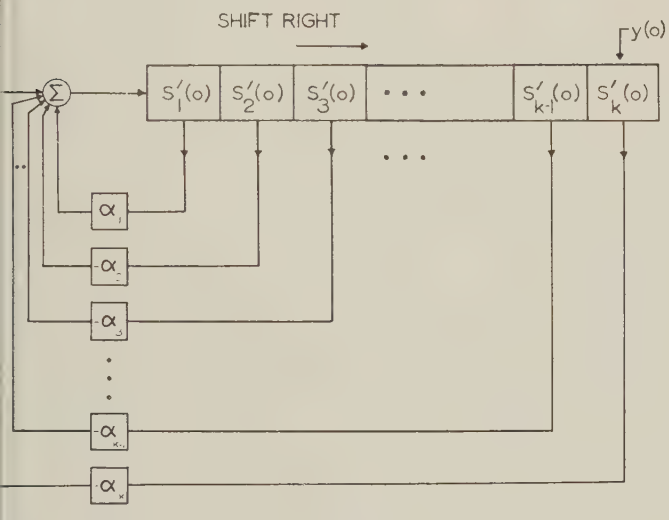


Fig. 4—Unexcited LMSC in "virtual" initial state.

Note that the unexcited output of the LMSC is simply the number which appears at the right-hand end of the register. Note, also, that in the unexcited case, since the register is assumed to shift right at each clock pulse,

$$s_i(n) = s_{i-1}(n-1) (i = 2, 3, \cdots, k). \tag{14}$$

It can be deduced from Fig. 4, or equivalently from (13) and (14), that each state of the unexcited LMSC appears in the form of a subsequence of length k imbedded in the impulse response. For example, the impulse response $h^{-1}(n)$ used in example B contains each successive state of the LMSC, $H^{-1}(z)$, as a subsequence of length 3. This structure is placed in evidence in Table II, where a portion of $h^{-1}(n)$ is displayed together with the succession of states of the LMSC which produces it.

TABLE II
Impulse Response and State Sequence

n	0	1	2	3	4	5	6	7	8	9
$h^{-1}(n)$	1	4	0	3	0	2	4	4	1	3
$s_1(n)$	0	3	0	2	4	4	1	3	4	0
$s(n)$ $s_2(n)$	4	0	3	0	2	4	4	1	3	4
$s_3(n)$	1	4	0	3	0	2	4	4	1	3
	$\mathbf{s}'(0)$									

VI. Generation of Periodic Sequences: Galois Theory

A critical examination of the procedure outlined in section IV leads immediately to several fundamental questions: 1) How can we be assured of the existence of a network which will produce an impulse response sequence of the desired periodicity? 2) How do we know that the sequence will have the required internal structure? 3) Is there a systematic method for constructing such a network? Fortunately, the mathematical theory of Galois fields answers the first two of these questions in the affirmative, and suggests an answer to the third. A summary of the portions of Galois theory which are pertinent to this discussion is presented in the appendix.

It has been demonstrated in Friedland and Stern[2] that an isomorphism exists between the set of all powers of any square matrix over the modular field, $GF(p)$ with an irreducible minimum polynomial[10] and the elements of a multiplication group of a Galois field. We make use of this isomorphism to construct periodic sequences. In an earlier paper,[1] it has been shown that if the characteristic polynomial of a matrix A over $GF(p)$ is irreducible, then A is periodic with period T, (*i.e.* $A^T = I$, and T is the smallest exponent for which this is true). It has also been shown that T is a divisor of $p^k - 1$. If a matrix of this type is the characteristic matrix of an LMSC, then the state sequence of the unexcited circuit, starting in any non-zero initial state $\mathbf{s}'(0)$ is periodic with the same period; *i.e.*, $\mathbf{s}'(T) = \mathbf{s}'(0)$. Consequently, the impulse response is also periodic with period T. It has further been shown that there exist matrices of this type with periods of maximal length $T = p^k - 1$; and hence it is possible to synthesize an LMSC whose impulse response has the period $p^k - 1$ [see (13)]. It follows that every non-zero state (or, equivalently, every non-zero subsequence of length k, in $h(n)$) must appear once and only once in each period of such a sequence. But this is exactly the type of structure that is necessary to fulfill conditions (9) in section IV. (It will become apparent as we proceed further, that all maximal period sequences have the required property that they contain $p - 1$ disjoint and distinct subsequences of length l, each one of which is a multiple of the first.)

[10] The minimum polynomial $m'(x)$ of a matrix A is the polynomial of lowest degree such that $m'(A) = 0$. It is either identical to the characteristic polynomial or a factor thereof. In the latter case, A is said to be derogatory. (S. Perlis, "Theory of Matrices," Addison-Wesley Press, Cambridge, Mass., p. 147; 1952.)

The decoding problem has now been reduced to one of synthesizing an A matrix which is periodic with maximum length period. Let us consider a $k \times k$ matrix A with characteristic equation $m(\lambda)$ irreducible over $GF(p)$. The isomorphism between powers A^i of the matrix, and the elements $p_i(\lambda)$ of the Galois group of $GF[p, m(\lambda)]$ is placed in evidence by expressing each power of A as a matrix polynomial in A of order $< k$. (That this is always possible, is a direct consequence of the Cayley-Hamilton theorem: a matrix satisfies its own characteristic equation.) Then, each polynomial in A is isomorphic to the same polynomial in λ. In order that A have a period of maximal length, its corresponding element, λ, must be a primitive element of the group of $GF[p, m(\lambda)]$ (see Appendix). We will now demonstrate that by proper selection of the polynomial, $m(\lambda)$, it is always possible to make λ a primitive element, and therefore make the period of A maximal length.

Consider the expression

$$f_n(x) = \frac{(x^n - 1) \cdot \prod (x^{n/p_i p_j} - 1) \cdot \prod (x^{n/p_i p_j p_k p_l} - 1) \cdots}{\prod (x^{n/p_i} - 1) \cdot \prod (x^{n/p_i p_j p_k} - 1) \cdots} \tag{15}$$

where $p_1, p_2, \cdots p_i$ are the distinct prime factors of n, and $n = p^k - 1$, p a prime. The products, $\prod$, are taken for all the combinations of distinct p's in the numbers indicated, each product $\prod$ in the numerator referring to an even number of the p's and each one in the denominator referring to an odd number of p's.

It can be shown[11] that $f_n(\lambda)$ is a polynomial of degree degree $\varphi(n)$ [where $\varphi(n)$ = the number of integers less than n, relatively prime to n] with integral coefficients, and that it contains all of the primitive n'th roots of unity. [The elements of the Galois group of $GF(p^k)$ are isomorphic to the $(p^k - 1)$th roots of unity.] The polynomial, $f_n(\lambda)$ is basic to our problem since it can be factored, mod p, into the form

$$f_n(\lambda) = q_1(\lambda) q_2(\lambda) \cdots q_s(\lambda) \mod p$$

where each $q_i(\lambda)$ is a monic k'th order polynomial irreducible over $GF(p)$. In addition, it can be shown[11] that the necessary and sufficient condition for an element α of $GF[p, m(\lambda)]$ to be primitive is that it satisfy the expression

$$q_i(\alpha) = 0 \mod m(\lambda)$$

for some $q_i(\lambda)$. But if we choose our modulus polynomial $m(\lambda)$ to be identical with some $q_i(\lambda)$ then λ itself satisfies that factor mod $m(\lambda)$. Thus, we deduce the following theorem.

Theorem: Given a $k \times k$ matrix A over the modular field $GF(p)$, the necessary and sufficient condition for A to be periodic with maximal length period is that its characteristic equation $m(\lambda)$ be a factor, mod p, of $f_{p^k-1}(\lambda)$.

This theorem can be clarified by an example.

Example C: Design of Filter of Example I

We have $p = 5$, $k = 2$.

In this case, the maximum length period is $T = 5^2 - 1 = 24$. First, form $f_{24}(\lambda)$.

$$f_{24}(\lambda) = \lambda^8 + 4\lambda^4 + 1$$

$\phi(24) = 8$, so that f_{24} is of eighth degree. It can be factored into four quadratics, each of which is irreducible over modular field $GF(p)$

$$f_{24}(\lambda) = (\lambda^2 + 4\lambda + 2)(\lambda^2 + 3\lambda + 3)(\lambda^2 + \lambda + 2)(\lambda^2 + 2\lambda + 3).$$

We may now construct an A matrix in companion form, whose characteristic equation is any one of the factors, say $m(\lambda) = \lambda^2 + 4\lambda + 2$.

$$A = \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix}.$$

Table III lists all powers of A with their equivalent polynomial forms. As the table indicates, A does indeed have a period of length 24. We may also construct a transfer function based upon $m(\lambda)$

$$H(z) = \frac{z^2}{z^2 + 4z + 2}.$$

Note that the donimator is derived from $m(\lambda)$, but the numerator can be chosen more or less arbitrarily, each choice resulting only in a cyclic permutation of terms of the basic impulse response. From $H(z)$, we can calculate $h(n)$ and observe that it is of maximal length period as expected.

$$h(n) = 1\ 1\ 4\ 2\ 4\ 0, 2\ 2\ 3\ 4\ 3\ 0, 4\ 4\ 1\ 3\ 1\ 0, 3\ 3\ 2\ 1\ 2\ 0; 1\ 1\ 4 \cdots$$

It should be noted that all higher powers of A which are relatively prime to 24, *i.e.*, 5, 7, 11, etc., are also primitive elements. Therefore, any one of them can also be used as the generator of a maximal length sequence. However, the higher powers of A will not in general be in the companion form used in (12). It is, therefore, necessary to reduce them to companion form by means of a similarity transformation. (A similarity transformation leaves the characteristic equation unchanged.)

For example,

$$A^7 = \begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix} \text{ is similar to } \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}.$$

[11] R. D. Carmichael, "Introduction to the Theory of Groups of Finite Order," Dover Publications, Inc., New York, N. Y., pp. 242–288; 1956.

TABLE III

Power	Matrix	Polynomial	Power	Matrix	Polynomial
A^1	$\begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix}$	$A^1 = A$	A^{13}	$\begin{bmatrix} 4 & 2 \\ 4 & 0 \end{bmatrix}$	$4A$
A^2	$\begin{bmatrix} 4 & 3 \\ 1 & 3 \end{bmatrix}$	$A^2 = A + 3I$	A^{14}	$\begin{bmatrix} 1 & 2 \\ 4 & 2 \end{bmatrix}$	$4A + 2I$
A^3	$\begin{bmatrix} 2 & 2 \\ 4 & 3 \end{bmatrix}$	$A^3 = 4A + 3I$	A^{15}	$\begin{bmatrix} 3 & 3 \\ 1 & 2 \end{bmatrix}$	$A + 2I$
A^4	$\begin{bmatrix} 4 & 1 \\ 2 & 2 \end{bmatrix}$	$A^4 = 2A + 2I$	A^{16}	$\begin{bmatrix} 1 & 4 \\ 3 & 3 \end{bmatrix}$	$3A + 3I$
A^5	$\begin{bmatrix} 0 & 2 \\ 4 & 1 \end{bmatrix}$	$A^5 = 4A + I$	A^{17}	$\begin{bmatrix} 0 & 3 \\ 1 & 4 \end{bmatrix}$	$A + 4I$
A^6	$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$	$A^6 = 2I$	A^{18}	$\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$	$3I$
A^7	$\begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix}$	$A^7 = 2A$	A^{19}	$\begin{bmatrix} 3 & 4 \\ 3 & 0 \end{bmatrix}$	$3A$
A^8	$\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$	$A^8 = 2A + I$	A^{20}	$\begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix}$	$3A + 4I$
A^9	$\begin{bmatrix} 4 & 4 \\ 3 & 1 \end{bmatrix}$	$A^9 = 3A + I$	A^{21}	$\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$	$2A + 4I$
A^{10}	$\begin{bmatrix} 3 & 2 \\ 4 & 4 \end{bmatrix}$	$A^{10} = 4A + 4I$	A^{22}	$\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}$	$A + I$
A^{11}	$\begin{bmatrix} 0 & 4 \\ 3 & 2 \end{bmatrix}$	$A^{11} = 3A + 2I$	A^{23}	$\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}$	$2A + 3I$
A^{12}	$\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$	$A^{12} = 4I$	A^{24}	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	I

Note that this matrix has a characteristic equation $\lambda^2 + 3\lambda + 3 = 0$, which is a factor of f_{24}, thus confirming the fact that it has a maximum length period. From this example, it can be observed that once one matrix of maximal length period has been determined, all other matrices of maximal length period are easily calculated.

Sometimes it is difficult to perform the factorization indicated in Example C. This is especially true for $f_n(\lambda)$ of high degree. However, the example still offers an indication of how to proceed. For example, consider the case illustrated in example B: $T_M = 124$, and $\phi(124) = 60$. Therefore, there exist 20 monic polynomials of third degree which are appropriate characteristic equations. Now, out of the 125 monic third order polynomials over $GF(5)$, 85 are reducible. This leaves only 40 irreducible polynomials left from which to choose. Thus, by selecting any irreducible polynomial at random we have a probability of 0.5 of finding an appropriate one.

VII. Small Error Correction

The foregoing material has been principally devoted to the case of unrestricted single errors; *i.e.*, $w = p - 1$. However, extension to other types of single errors is quite straightforward. If we are to consider only close-packed codes, (2) with the equals sign must have an integer solution. It has already been stated that such is the case for $w = p - 1$. It can also be shown that this is always the case for $w = 1, 2$; with the understanding that $w \leq p - 1$ in all cases. (It should be noted here that when we restrict ourselves to smaller magnitudes of errors, the allowable word length l for a given number of check digits increases.) For other values of w it may not always be possible to construct a close-packed code. For example, the case $p = 7$, $k = 3$, $w = 4$ does not have an integer solution for l. Therefore, to maintain some sort of generality, we will consider only the cases $w = 1, 2$. The former case might correspond to a channel in which transition from one level to the next lowest is the only type of error which is likely. In this case, an error would correspond to an impulse of magnitude $p - 1$. The latter case might correspond to the small error case considered by Ulrich,[8] where each digit may change to the level just above or just below. (These models may be somewhat unrealistic in that a transition from zero to $p - 1$ is possible in the model, although it would probably be unlikely in practice.) Let us consider these two cases individually.

$w = 1$: In this case, $l = p^k - 1$; *i.e.*, the block length is just equal to the period of the maximal length sequence. Assume that the magnitude of the expected error is m. Then, an error in the i'th place of the block will produce a response, $mh^{-1}(n - i - 1)$, where $h^{-1}(n)$ is the impulse response of the decoding filter. Thus, we must now compare the last k digits to mh^{-1} rather than h^{-1}. With this modification, the decoding procedure is the same as that described in section IV.

$w = 2$: In this case, the block length is just one half of the maximal length sequence. An important consideration here is that an $h^{-1}(n)$ must be chosen which prevents ambiguity. For example, consider the case $p = 5$, $k = 3$, and error magnitudes of either 1 or $p - 1$, (one level up or one level down). In this case, $l = 62$, which means that the block length covers two of the four subsequences of h^{-1}. If one of these two subsequences happened to be $p - 1$ times the other, it would be impossible to distinguish between an error of magnitude 1 occurring in one half of the block an an error of magnitude $p - 1$ occurring in the other half. Fortunately, it is generally true that any $k \times k$ matrix possessing a maximal length period has the property that

$$A^{p^{k-1/2}} = (p - 1)I \qquad (p > 2).$$

This assures that there will be no ambiguity in the case just described. For other types of small errors, however, it is necessary to select an unambiguous sequence.

VIII. Construction of a Coding and Decoding System

To complete this discussion, a physical system which realizes the aforementioned coding and decoding systems is presented. We shall consider only the case of unrestricted errors. (Small errors can be handled with minor modifications.) The coding system is simply an LMSC having a transfer function $H(z)$.

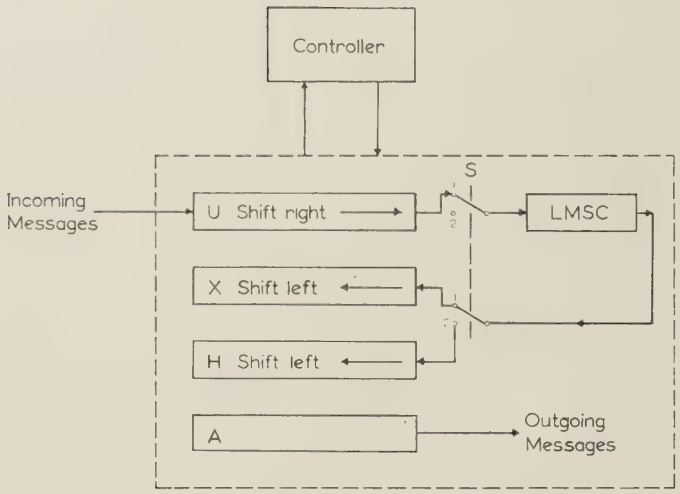


Fig. 5—Decoding machine.

Fig. 5 shows the structure of the decoding system. Basically, it is a small computer comprising a controller, three shift registers, an LMSC and an accumulator. The function of each component is described below:

1) Controller: Contains the program and the logic elements. Controls the operation of the decoder.

2) U Register: A shift register which stores the incoming message, $u'(n)$ from right to left. During decoding, u' is shifted to the *right* into LMSC.

3) LMSC: A modular sequential circuit having transfer function, $H^{-1}(z)$. Its operation, and that of all shift registers, is initiated by a shift pulse, so that the LMSC advances one state with each pulse.

4) X Register: A shift register which stores $x'(n)$, the output of the decoding filter, H^{-1}. When the switch S is in position 1, this register is loaded from the right, by the LMSC.

5) H Register: A shift register which is loaded from the right by the LMSC when the switch S is in position 2. Its function is to store portions of the impulse response $h^{-1}(n)$ for comparison with the check digits of $x'(n)$.

6) A Register: An accumulator used for comparing check digits and for storing the corrected message $x(n)$. (All registers are of length l.)

A typical program for this machine might be as follows.

Program

1) Reset LMSC and registers X, H, and A to zero.
2) Set S to position 1; read contents of U through LMSC into X (this requires l shift pulses).
3) Set S to position 2; reset LMSC to initial state, $\mathbf{s}'(0)$.
4) Add contents of X to A.
5) Add complement of contents of H to A.
6) Examine last k digits of A:
   a) If they are zero go to instruction 11;
   b) If they are not zero, go to instruction 7.
7) Reset A to zero.
8) Examine the first digit of H:
   a) If it is zero, go to instruction 9;
   b) If it is not zero go to instruction 13.
9) Apply one shift pulse to LMSC.
10) Go to instruction 4.
11) Print out contents of A.
12) Go to instruction 1.
13) Reset H to zero.
14) Go to instruction 9.

Appendix

In the present appendix are summarized some of the definitions and properties of a Galois field of order p^k (which we shall alternatively call a field of *k-dimensions*). For proofs of the properties listed below the reader is referred to a standard text on group theory.[11]

Elements of $GF(p^k)$

The set of all polynomials of degree $k-1$ with coefficients in $GF(p)$ constitutes the elements of a k-dimensional Galois field $GF(p^k)$. Specifically each element is of the form

$$\alpha = a_k + a_{k-1}\lambda + \cdots + a_1\lambda^{k-1}$$

where $a_1, a_2, \cdots, a_k$ range over the integers $0, 1, \cdots, p-1$. There are clearly p^k elements in this field.

Addition

Addition in $GF(p^k)$ is defined as the sum, modulo p, of pairs of polynomials; *i.e.*, if

$$\alpha = a_k + a_{k-1}\lambda + \cdots + a_1\lambda^{k-1}$$

and

$$\beta = b_k + b_{k-1}\lambda + \cdots + b_1\lambda^{k-1}$$

then

$$r = \alpha + \beta = (a_k + b_k) + (a_{k-1} + b_{k-1})\lambda + \cdots + (a_1 + b_1)\lambda^{k-1} \mod p.$$

Clearly the set is closed under addition, and the following properties are present:

$$\alpha + \beta = \beta + \alpha$$

$$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma.$$

There exists an x in $GF(p^k)$ for every α such that

$$\alpha + x = 0 = 0 + 0\lambda + \cdots + 0\lambda^{k-1}.$$

Multiplication

Multiplication in the Galois field $GF(p^k)$ is defined as the remainder obtained by dividing the product modulo p of a pair of elements by an irreducible *base polynomial* in $GF(p)$ of degree k, *i.e.*,

$$\alpha \cdot \beta \equiv \alpha\beta = (a_k + a_{k-1} + \cdots + a_1\lambda^{k-1}) \cdot (b_k + b_{k-1}\lambda + \cdots + b_1\lambda^{k-1}) \mod [p, m(\lambda)].$$

The double modulus, notation mod $[p, m(\lambda)]$ indicates that the product of the elements mod p is first formed, and then the remainder upon dividing this product by $m(\lambda)$ is found. The product depends on the irreducible polynomial; hence, the field may be explicitly identified by using the notation $GF[p, m(\lambda)]$. Clearly, the set of elements is closed under this operation, and the following properties obtain:

$$\alpha \cdot \beta = \beta \cdot \alpha$$

$$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma.$$

There exists an x in $GF(p^k) = GF[p, m(\lambda)]$ for every $\alpha \neq 0$ such that

$$\alpha x = 1 = 1 + 0\lambda + \cdots + 0\lambda^{k-1}.$$

In addition the distributive property holds

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma.$$

Powers of an Element

The non-zero elements of $GF(p^k)$ constitute an Abelian group under multiplication; hence for every $\alpha \pm 0$, there exists a μ such that

$$\alpha^{\mu} = 1.$$

The smallest number $\mu = t$ for which the above equation is valid is called the *period*[12] of α [which depends on $m(\lambda)$]. The set of elements

$$\alpha, \alpha^2, \cdots, \alpha^{t-1}, \qquad \alpha^t = 1$$

constitute an Abelian subgroup of period t, with α as its generator.

The following theorems are of particular importance in connection with modular sequential circuits.

Therorem i: The period t of any subgroup of the multiplication group of $GF(p^k)$ is a divisor of $p^k - 1$.

Theorem ii: The number of elements of period t in $GF(p^k)$ is $\varphi(t)$, where $\varphi(t)$ is Euler's φ-function and is equal to the number of numbers ≥ 0 and $< t$ relatively prime to t.

If the period of an element is $p^k - 1$, the element generates the entire multiplication group of $GF(p^k)$. Such an element is called primitive. From Theorem ii, the number of primitive elements is $\varphi(p^k - 1)$.

The period of any particular element depends on the base polynomial which defines multiplication in the field. By the appropriate choice of a base polynomial, it is possible to make any element in the field except zero have any of the allowable periods.

[12] The term *period*, rather than *order* is used to avoid confusion with the order of an LMSC, although the latter is used in the mathematical literature.

Some Spectral Properties of Weighted Random Processes*

H. S. SHAPIRO† AND R. A. SILVERMAN†

Summary—**We study the power spectrum and, more generally, the spectral covariance of weighted stationary processes. It is found that if the power spectrum of the underlying stationary process is suitably well behaved and properly matched to the weight function, then the high-frequency behavior of the power spectrum and spectral covariance is especially simple. Asymptotic theorems describing this behavior precisely are given.**

Introduction

IN many problems involving random processes, the object of primary interest is a stationary random process, $x(t)$ say, but what is actually available for observation instead is the weighted process $y(t) = m(t)x(t)$, obtained by multiplying $x(t)$ by the nonrandom weight function $m(t)$. For example, in analyzing a radio scattering experiment, the dielectric noise $x(\mathbf{r})$,[1] which is usually regarded as spatially stationary, is multiplied by a weight function $m(\mathbf{r})$ representing the joint effect of the gain patterns of the transmitting and receiving antennas, and only the region where $m(\mathbf{r})$ is large, the so-called scattering volume, makes an appreciable contribution to the received scattered signal.[2] Further examples abound in the important problem of experimental power spectrum measurement, where the weight functions $m(t)$ correspond to the *data windows* discussed by Blackman and Tukey;[3] these

* Manuscript received by the PGIT, March 13, 1959. The research reported in this article has been sponsored by the AF Cambridge Res. Center, Air Res. and Dev. Command, under Contract No. AF 19(604)3495, and by the Office of Naval Res., under Contract No. N6ori-201(01).

† Inst. of Math. Sciences, N. Y. U., 25 Waverly Pl., New York 3, N. Y.

[1] By the dielectric noise we mean the refractive index fluctuations usually attributed to the action of atmospheric turbulence. Here we replace the scalar t (time) by the vector $\mathbf{r}$ (position), since we are dealing with a random field.

[2] PROC. IRE, Scatter Propagation Issue, vol. 43; October, 1955.

[3] R. B. Blackman and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering," *Bell Sys. Tech. J.*, vol. 37, pp. 185–282, January, 1958; pp. 485–569, March, 1958.

authors give a detailed analysis of the practical consequences of various choices of $m(t)$. Weighted processes have also been studied by Pugachev[4] and Burford.[5] Indeed, in a sense, weighted processes are more physical than stationary processes, since the latter have the unreasonable property of continuing forever with constant average noise power. (Of course, this property is never taken too seriously by applied people, who customarily allude to stationary processes as convenient idealizations of physical noises.)

The question with which we are mainly concerned in this paper is roughly the following: Under what conditions will the power spectra of $x(t)$ and $y(t)$ be approximately the same for sufficiently high frequencies? [For, whereas it is to be expected that in general $x(t)$ and $y(t)$ will have quite different power spectra at low frequencies, because of the modulatory action of $m(t)$, it is quite possible for the spectra to be nearly the same at sufficiently high frequencies.] We find that to give even a partial answer to this question we must explore the asymptotic behavior of convolutions.

Preliminary Analysis

Let $x(t)$ be a zero-mean stationary random process, so that $Ex(t) = 0$, $Ex(t)\bar{x}(t') = C(t - t')$, where, as usual, E denotes the expectation or ensemble average, and the overbar denotes the complex conjugate. We assume that the autocorrelation function $C(\tau)$ is continuous and absolutely integrable, and write the Wiener-Khintchine relations in the form

$$C(\tau) = \int_{-\infty}^{\infty} \exp(i\omega\tau) f(\omega)\, d\omega,$$

$$f(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\omega\tau) C(\tau)\, d\tau,$$

where the power spectrum (more accurately, the power spectral density) $f(\omega)$ is non-negative, continuous and integrable; if $x(t)$ is real, $f(\omega)$ is even, but more generally we allow $x(t)$ to be complex. Now let $m(t)$ be a bounded continuous function, and construct the weighted process $y(t) = m(t)x(t)$. Clearly $Ey(t) = 0$, $Ey(t)\bar{y}(t') = m(t)\bar{m}(t')C(t - t')$, so that in particular $y(t)$ is not stationary. [However, the normalized process $y(t)/(E|y(t)|^2)^{1/2}$ is stationary.] We assume that $m(t)$ is square-integrable, which implies that the sample functions of $y(t)$, unlike those of $x(t)$, are themselves square-integrable with probability one. To see this, note that

$$E \int_{-\infty}^{\infty} |\, y(t)\,|^2\, dt = \int_{-\infty}^{\infty} E\,|\, y(t)\,|^2\, dt$$

$$= C(0) \int_{-\infty}^{\infty} |\, m(t)\,|^2\, dt < \infty.$$

We can therefore take L^2 Fourier transforms of the sample functions of $y(t)$ individually, obtaining the spectral process[6,7]

$$Y(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\omega t) y(t)\, dt \tag{1}$$

with (spectral) covariance $\Gamma(\omega, \omega') = EY(\omega)\overline{Y}(\omega')$.

We now find the relation between $\Gamma(\omega, \omega')$ and the power spectrum of $x(t)$. First we observe that

$$x(t) = \int_{-\infty}^{\infty} \exp(i\omega t)\, dX(\omega), \tag{2}$$

where the integral is meant in the mean square sense, and the increments of the spectral process $X(\omega)$ obey the symbolic relation[8,9]

$$E\, dX(\omega)\, \overline{dX}(\omega') = f(\omega)\, \delta(\omega - \omega')\, d\omega\, d\omega'. \tag{3}$$

It follows from (1) and (2) that

$$Y(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\omega t) m(t) x(t)\, dt$$

$$= \int_{-\infty}^{\infty} M(\omega - \omega')\, dX(\omega'),$$

where

$$M(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\omega t) m(t)\, dt$$

is the Fourier transform of the weight function $m(t)$. By Plancherel's theorem, $M(\omega)$ is itself square-integrable, and in fact $\int_{-\infty}^{\infty} |M(\omega)|^2\, d\omega = (1/2\pi) \int_{-\infty}^{\infty} |m(t)|^2 dt$. Finally, forming the covariance of $Y(\omega)$ and using (3), we obtain

$$\Gamma(\omega, \omega') = \int_{-\infty}^{\infty} M(\omega - s)\overline{M}(\omega' - s) f(s)\, ds, \tag{4}$$

the desired relation between $\Gamma(\omega, \omega')$ and $f(\omega)$. The quantity $\Gamma(\omega, \omega) = E\,|Y(\omega)|^2$ is appropriately called the power spectrum of $y(t)$, even though $y(t)$ is not stationary; for simplicity we shall henceforth abbreviate $\Gamma(\omega, \omega)$ to just $\gamma(\omega)$. Specializing (4) to the case $\omega = \omega'$, we obtain

$$\gamma(\omega) = \int_{-\infty}^{\infty} K(\omega - s) f(s)\, ds = \int_{-\infty}^{\infty} K(s) f(\omega - s)\, ds, \tag{5}$$

where we have introduced the notation $K(\omega) \equiv |M(\omega)|^2$. Thus, the power spectrum of $y(t)$ is the convolution of the power spectrum of $x(t)$ with the non-negative integrable kernel $K(\omega)$, a familiar result.[3] In what follows, we shall assume, as we may without loss of generality, that $\int_{-\infty}^{\infty} K(\omega) d\omega = 1$; otherwise, we need only supply an appropriate constant factor in certain obvious places.

Asymptotic Behavior of $\gamma(\omega)$

Examining the convolution (5), we expect that if the

[4] V. S. Pugachev, "The Theory of Random Functions and its Application to Problems of Automatic Control," Gos. Izdat. Tekh.-Teor. Lit., Moscow; 1957. (In Russian.)

[5] T. M. Burford, "Nonstationary velocity estimation," *Bell Sys. Tech. J.*, vol. 37, pp. 1009–1021; July, 1958.

[6] By $f(\omega) \in L^p$, we mean, as usual, that the (Lebesgue) integral $\int_{-\infty}^{\infty} |f(\omega)|^p d\omega$ is finite.

[7] Here, and occasionally below, the integral means l.i.m.$_{A\to\infty} \int_{-A}^{A}$.

[8] $\delta(\omega)$ is the Dirac delta function.

[9] M. Loève, "Probability Theory," D. Van Nostrand, Inc., New York, N. Y., ch. 10; 1955.

ernel $K(\omega)$ is small compared to $f(\omega)$ as $\omega \to \infty$, and if $\omega)$ does not vary too rapidly, then $\gamma(\omega)$ should behave ke $f(\omega)$ for large ω, since $\gamma(\omega)$ is an average of $f(s)$ with eight $K(\omega - s)$ centered at $s = \omega$. [If, on the other hand, (ω) is *large* compared to $f(\omega)$, the roles of $f(\omega)$ and $K(\omega)$ e interchanged.] Speaking even more qualitatively, hen two functions are convolved, it should often be ossible to regard the function which falls off more rapidly s averaging the function which falls off less rapidly. We hall see that this intuitive idea can be made precise under irly general hypotheses, but that it breaks down when $\omega)$ is too rapidly decreasing. First we must specify what e mean by a function which does not vary too rapidly.

Definition. A function $f(\omega)$ is slowly varying as $\omega \to \infty$ riefly, slowly varying) if $f(\omega)/f(\sigma)$ tends uniformly to 1 henever $\omega, \sigma \to \infty$ in such a way that $\omega/\sigma \to 1$.

Slowly varying functions are related to the *slowly scillating* functions that arise in connection with auberian theorems. A slowly oscillating function $g(\omega)$ one for which $|g(\omega) - g(\sigma)| \to 0$ uniformly whenever $\sigma \to \infty$ in such a way that $\omega/\sigma \to 1$;[10] a slowly varying nction, as defined above, is just the exponential of a owly oscillating function. For example, the functions $+ |\omega|^p)^{-1}$ and $(1 + |\omega|^p)^{-1} \log (1 + |\omega|)$, $p > 0$ are owly varying, but $\exp(-|\omega|^p)$, $p > 0$ is not. The follow- ng facts about slowly varying functions are an immediate nsequence of the definition.

1) If $f(\omega)$ is slowly varying, it does not vanish for fficiently large ω; this observation is important in con- ection with certain expressions written below where $\omega)$ appears in the denominator.

2) The product and quotient of two slowly varying nctions are slowly varying.

3) If for large ω the function $g(\omega)$ has a derivative and $|g'(\omega)| \leq C/\omega$, then $|g(\omega) - g(\sigma)| = |\int_\sigma^\omega g'(s)ds| \leq ((\omega/\sigma) - 1)$, so that $g(\omega)$ is slowly oscillating. Thus, for $\omega)$ to be slowly varying, it is sufficient that $f(\omega)$ be fferentiable and nonvanishing and that $\omega(d/d\omega) \log f(\omega)$ e bounded, all for sufficiently large ω.

4) If $f(\omega)$ is slowly varying, positive and continuous, hen if $q > 0$ is sufficiently large, we have $f(\omega) < \omega^q$ for rge ω. The proof is as follows: For $\omega > \omega_0$, $f(\omega)/f(\rho\omega) \leq 2$, ρ is sufficiently close to 1, $0 < \rho < 1$. Iterating this equality n times, where n is the largest integer such hat $\rho^{n-1}\omega > \omega_0$, we obtain $f(\omega) \leq 2^n f(\rho^n\omega) \leq M2^n$, where I is the maximum of $f(\omega)$ in the interval $(\rho\omega_0, \omega_0)$. Now $^n = \rho^{-cn} < (\rho\omega_0/\omega)^{-c} = A\omega^c$, where $\rho^{-c} = 2$, so that $\omega) \leq MA\omega^c$, which proves the result. Moreover, since $/f(\omega)$ is slowly varying, we also have $f(\omega) > \omega^{-p}$ for itable p and large ω.

It follows from the last remark that slowly varying nctions are, roughly speaking, functions which behave ke a power of ω at infinity. It will be noted that the lass of slowly varying power spectra is comfortably rge: for example, it contains all spectra obtained by passing white noise through linear networks, but not band-limited spectra. (However, band-limited spectra are physically unrealizable.)

Returning to our heuristic notions, we require that the kernel $K(\omega)$ be small compared to $f(\omega)$ for large ω. The mathematical condition which expresses this most naturally is that

$$\lim_{\omega\to\infty} \frac{K(\omega)}{f(\omega)} = 0. \tag{6}$$

It is important to note that (6) implies nothing about the relative "widths" of $f(\omega)$ and $K(\omega)$. For example, the "halfwidth" of $K(\omega) = T/\pi(1 + (T\omega)^2)$ is $1/T$, while that of $f(\omega) = 1/(1 + |\omega|^p)$, $0 < p < 2$, is 1, but with this choice (6) is valid for *any* $T > 0$. Thus, no restriction on the width of data windows will figure in our results.

We are now in a position to state Theorem 1.

Theorem 1: (*Asymptotic convolution theorem*) *Let $f(\omega)$ and $K(\omega)$ be non-negative integrable functions (with $\int_{-\infty}^{\infty} K(\omega)d\omega = 1$). If $f(\omega)$ is slowly varying and non-increasing for sufficiently large ω, and if*

$$\lim_{\omega\to\infty} \frac{K(\omega)}{f(\omega)} = 0, \tag{6}$$

then

$$\lim_{\omega\to\infty} \frac{\gamma(\omega)}{f(\omega)} = 1, \tag{7}$$

where $\gamma(\omega)$ is the convolution (5).

Remark 1: Of course, the theorem remains true if ∞ is replaced by $-\infty$ in the hypotheses and conclusions.

Remark 2: Actually, we need not assume that $f(\omega)$ is nonincreasing, but the milder restriction

$$f(\sigma) \leq Cf(\omega), \quad \text{when} \quad \omega_0 < \omega < \sigma, \tag{8}$$

where C is a constant independent of ω.

Proof of Theorem 1: We have

$$\gamma(\omega) = \int_{-\infty}^{R} K(s)f(\omega - s)\, ds + \int_{R}^{\infty} K(s)f(\omega - s)\, ds, \tag{9}$$

where $R = R(\omega)$ is a function of ω to be specified later. For the moment we suppose only that

$$R(\omega) \to \infty \quad \text{as} \quad \omega \to \infty, \tag{10}$$

$$\lim_{\omega\to\infty} \frac{R(\omega)}{\omega} = 0. \tag{11}$$

Let us denote the two integrals in (9) by $\gamma_1(\omega)$ and $\gamma_2(\omega)$, respectively. We consider first the quantity

$$\gamma_1(\omega)/f(\omega) = \int_{-\infty}^{R} K(s)(f(\omega - s)/f(\omega))\, ds,$$

which it is convenient to rewrite as

$$\gamma_1(\omega)/f(\omega) = \int_{-\infty}^{\infty} K(s)\phi_R(s)(f(\omega - s)/f(\omega))\, ds, \tag{12}$$

where $\phi_R(s)$ is the characteristic function of the interval $(-\infty, R(\omega))$, *i.e.*, the function equal to 1 when s lies

[10] G. H. Hardy, "Divergent Series," Clarendon Press, Oxford, ng.; 1949. See p. 124.

in $(-\infty, R(\omega))$ and to 0 otherwise. Now the ratio $f(\omega - s)/f(\omega)$ is bounded in the two-dimensional region $\omega > \omega_0$, $-\infty < s < R(\omega)$, where ω_0 is a suitably large number. For $0 \leq s < R(\omega)$, this follows from (11) and from the fact that $f(\omega)$ is slowly varying; for $s < 0$, it follows from the fact that $f(\omega)$ is nonincreasing [or, more generally, satisfies (8)]. Moreover, for each fixed s, $\lim_{\omega\to\infty} \{f(\omega - s)/f(\omega)\} = 1$ [again because $f(\omega)$ is slowly varying]. Thus, since $K(\omega)$ is integrable, we can use the Lebesgue dominated convergence theorem to justify passing to the limit under the integral in (12), with the result

$$\lim_{\omega\to\infty} \frac{\gamma_1(\omega)}{f(\omega)} = \int_{-\infty}^{\infty} K(s)\, ds = 1.$$

Consequently, the proof of Theorem 1 will be complete when we show that

$$\lim_{\omega\to\infty} \frac{\gamma_2(\omega)}{f(\omega)} = 0. \tag{13}$$

Now, in view of (6), we can write

$$K(\omega) = \epsilon(\omega) f(\omega),$$

where $\epsilon(\omega) \to 0$ as $\omega \to 0$. This gives

$$\gamma_2(\omega) = \int_R^{\infty} \epsilon(s) f(s) f(\omega - s)\, ds$$

for the second integral in (9). Therefore, since $f(\omega)$ is nonincreasing [or satisfies (8)]

$$\gamma_2(\omega) \leq \epsilon_0(R) C f(R) \int_{-\infty}^{\infty} f(\omega - s)\, ds = C' \epsilon_0(R) f(R),$$

where C' is a constant and

$$\epsilon_0(\omega) = \sup_{\sigma \geq \omega} \epsilon(\sigma).$$

Note that $\epsilon_0(\omega)$ is nonincreasing and $\to 0$ as $\omega \to 0$. Hence, to establish (13) and thereby complete the proof, it suffices to construct a function $R(\omega) \to \infty$ which satisfies (11) and is such that

$$\lim_{\omega\to\infty} \frac{\epsilon_0(R) f(R)}{f(\omega)} = 0. \tag{14}$$

Since $f(\omega)$ is slowly varying, there exists a number ρ, $0 < \rho < 1$, such that

$$f(\rho\omega) \leq 2 f(\omega), \tag{15}$$

provided that ω is suitably large. Iteration of (15) leads to the relation

$$f(\rho^n \omega) \leq 2^n f(\omega), \tag{16}$$

which holds for all n, again provided that ω is in each case suitably large. Now, since $\epsilon_0(\omega) \to 0$, we can construct a sequence $\omega_1 < \omega_2 < \omega_3 < \cdots$ of positive numbers with the properties

$$\begin{aligned} \omega_n &> n\rho^{-n}, \\ \epsilon_0(\rho^n \omega_n) &< 3^{-n}, \end{aligned} \tag{17}$$

and also large enough to guarantee the validity of (16) for $\omega \geq \omega_n$. We now define

$$R(\omega) = \rho^n \omega, \qquad \omega_n \leq \omega < \omega_{n+1}.$$

Since $R(\omega) \geq \rho^n \omega_n > n$, for $\omega_n \leq \omega < \omega_{n+1}$, (10) is satisfied. Moreover, since $R(\omega)/\omega = \rho^n$ for $\omega_n \leq \omega < \omega_{n+1}$, (11) holds. Finally, using (16) and (17) we have

$$\epsilon_0(R) f(R)/f(\omega) \leq 2^n \epsilon_0(\rho^n \omega_n) < (\tfrac{2}{3})^n,$$

for $\omega_n \leq \omega < \omega_{n+1}$, which establishes (14) and completes the proof of Theorem 1.

Examples and Extensions

We now give examples which illustrate the meaning of Theorem 1. In some examples one or the other of the hypotheses of the theorem is violated, with the result that $\lim_{\omega\to\infty} \{\gamma(\omega)/f(\omega)\} \neq 1$. In other examples, the requirements on $f(\omega)$ are weakened but those on $K(\omega)$ are strengthened, with the result that $\lim_{\omega\to\infty} \{\gamma(\omega)/f(\omega)\} = 1$ remains valid.

1) An especially simple weight function is

$$\begin{aligned} m(t) &= \sqrt{\pi/T}, && |\, t \,| \leq T, \\ m(t) &= 0, && |\, t \,| > T, \end{aligned}$$

i.e., a "wide open" data window. The corresponding (Fejér) kernel is

$$K(\omega) = M^2(\omega) = \sin^2 T\omega/\pi T\omega^2.$$

If now $f(\omega) = 1/(1 + |\omega|^p)$, $1 < p < 2$, then according to Theorem 1, no matter how close p is to 2 and regardless of the size of T, the convolution (5) approximates $f(\omega)$ for sufficiently large ω. (How large ω must be for good approximation will depend, of course, on p and T.) This result would be difficult to verify by direct calculation.

2) Even when it can be done, it is usually very tedious to verify Theorem 1 directly. For example, let $K(\omega) = \sqrt{2}/\pi(1 + \omega^4)$, $f(\omega) = 1/(1 + \omega^2)$, which satisfy the hypotheses of Theorem 1. The convolution of these two functions can be evaluated by residues and is found to be

$$\begin{aligned} \gamma(\omega) = {} & (\omega^6 + 6\omega^4 + 2\omega^2)/(\omega^8 + 4\omega^6 + 8\omega^4 - 8\omega^2 + 4) \\ & + \sqrt{2}(\omega^2 - 6)/(\omega^6 + 4\omega^4 + 4\omega^2 + 16). \end{aligned}$$

We see at once that $\lim_{\omega\to\infty} \{\gamma(\omega)/f(\omega)\} = 1$.

3) As an example of a case where condition (6) is violated, choose $K(\omega)$ and $f(\omega)$ to be the *same* slowly varying function $1/\pi(1 + \omega^2)$. Then the convolution (5) is found to be $\gamma(\omega) = 2/\pi(\omega^2 + 4)$, so that $\lim_{\omega\to\infty} \{\gamma(\omega)/f(\omega)\} = 2$ instead of 1.

4) Choose $f(\omega) = e^{-a|\omega|}$, $a > 0$, which is not slowly varying, and let $K(\omega)$ be decreasing for large ω and such that $\int_{-\infty}^{\infty} e^{a\omega} K(\omega) d\omega = A < \infty$. Then the convolution (5) is

$$\gamma(\omega) = e^{-a\omega} \int_{-\infty}^{\omega} e^{as} K(s)\, ds + e^{a\omega} \int_{\omega}^{\infty} e^{-as} K(s)\, ds.$$

Since for large enough ω the second integral does not exceed

$$e^{a\omega} K(\omega) \int_{\omega}^{\infty} e^{-as}\, ds = K(\omega)/a,$$

nd since

$$\lim_{\omega\to\infty}\frac{K(\omega)}{f(\omega)} = 0,$$

follows that

$$\lim_{\omega\to\infty}\frac{\gamma(\omega)}{f(\omega)} = A,$$

here in general $A \neq 1$.

5) Choose $f(\omega) = |\omega|e^{-|\omega|}$ and $K(\omega) = \frac{1}{2}e^{-|\omega|}$. Although $\omega)$ is not slowly varying, (6) is satisfied. The convolution 5) is an elementary integral and is found to be

$$\gamma(\omega) = \tfrac{1}{4}e^{-\omega}(1 + \omega + \omega^2), \qquad \omega > 0,$$

o that

$$\lim_{\omega\to\infty}\frac{\gamma(\omega)}{f(\omega)} = \infty.$$

6) As a more complicated example, choose $K(\omega) = \exp(-|\omega|^p)$, $p > 0$,[11] and $f(\omega) = \exp(-|\omega|^q)$, where $< q < p/(p+1)$. Again $f(\omega)$ is not slowly varying, ut (6) holds, and it is in fact true that

$$\lim_{\omega\to\infty}\{\gamma(\omega)/f(\omega)\} = 1.$$

o show this, we note first that in the proof of Theorem 1 e can replace the hypothesis that $f(\omega)$ be slowly varying y the weaker condition

$$\begin{aligned} &f(\omega)/f(\sigma) \to 1 \text{ } \textit{uniformly as } \omega, \sigma \to \infty \\ &\textit{in such a way that } |\omega - \sigma| < R(\omega), \end{aligned} \tag{18}$$

rovided that we strengthen the requirements on $K(\omega)$ to

$K(\omega)$ *is nonincreasing for large enough* ω,

$$\lim_{\omega\to\infty}\frac{K(R)}{f(\omega)} = 0. \tag{19}$$

or then $\lim_{\omega\to\infty}\{\gamma_1(\omega)/f(\omega)\} = 1$, by an obvi- us extension of the first part of Theorem 1, and ince $\gamma_2(\omega) = \int_R^\infty K(s)f(\omega - s)ds$ is bounded above y $K(R)\int_{-\infty}^{\infty} f(\omega - s)ds$, $\lim_{\omega\to\infty}\{\gamma_2(\omega)/f(\omega)\} = 0$ by 19), so that $\lim_{\omega\to\infty}\{\gamma(\omega)/f(\omega)\} = 1$ is still valid. Now et $R(\omega) = \omega^\alpha$, $0 < \alpha < 1$, where α will be chosen later. hen, if $0 < \omega < \sigma$, we have

$$\begin{aligned} (\omega)/f(\sigma) &= \exp(\sigma^q - \omega^q) \\ &= \exp\{q(\sigma - \omega)\rho^{q-1}\}, \qquad \omega < \rho < \sigma, \end{aligned}$$

y the mean value theorem. If $\sigma - \omega < \omega^\alpha$, the exponent n the right will tend to zero [*i.e.*, (18) will hold] if $+ q < 1$. Moreover, $\lim_{\omega\to\infty}\{K(\omega^\alpha)/f(\omega)\} = 0$, pro- ided that we have $q < p\alpha$ as well as $\alpha + q < 1$. These nequalities are compatible if we can insert a number be- ween q/p and $1 - q$, *i.e.*, if $q/p < 1 - q$ or $q < p/(p+1)$, hich completes the demonstration of our example.

[11] The constant a is chosen to make

$$\int_{-\infty}^{\infty} K(\omega) = 1.$$

7) The conclusion of Theorem 1 remains true if we replace the hypothesis that $f(\omega)$ is integrable by the weaker hypothesis that $f(\omega)$ is *bounded*, provided that we also replace (6) by the stronger condition

$$\lim_{\omega\to\infty}\frac{\int_\omega^\infty K(s)\,ds}{f(\omega)} = 0.$$

Since integrability of $f(\omega)$ was used only in showing that $\gamma_2(\omega)$ satisfies $\lim_{\omega\to\infty}\{\gamma_2(\omega)/f(\omega)\} = 0$, it suffices to show that the same is true with the altered hypotheses. Let us define $\epsilon(\omega) = \int_\omega^\infty K(s)ds/f(\omega)$, which by hypothesis $\to 0$ as $\omega \to \infty$. Then, if $f(\omega) \leq M$,

$$\frac{\gamma_2(\omega)}{f(\omega)} \leq \frac{M\int_R^\infty K(s)\,ds}{f(\omega)} \leq \frac{M\epsilon(R)f(R)}{f(\omega)},$$

and as above we know that we can construct a function $R = R(\omega)$ satisfying (10) and (11) and such that $\lim_{\omega\to\infty}\{\epsilon(R)f(R)/f(\omega)\} = 0$. Hence the proof is complete. Applying this result to the kernel of example 1, we have $\int_\omega^\infty K(s)ds = O(1/\omega)$, so that the example is valid for $0 \leq p < 1$ as well as for $1 < p < 2$.

8) Continuing our discussion of example 1, we note that the case $p = 1$ is not covered by the arguments given so far. This case may be inferred from the following extension of Theorem 1: If $f(\omega)$ is assumed to be of class[6] $L^{q'}$, $q' > 1$, rather than of class L^1, and we define $q = q'/(q'-1)$, then the conclusion of Theorem 1 remains valid if (6) is replaced by the requirement

$$\lim_{\omega\to\infty}\frac{\left[\int_\omega^\infty [K(s)]^q\,ds\right]^{1/q}}{f(\omega)} = 0.$$

For we have only to apply Hölder's inequality, obtaining

$$\begin{aligned} \gamma_2(\omega) &= \int_R^\infty f(\omega - s)K(s)\,ds \\ &\leq \left[\int_R^\infty [f(\omega - s)]^{q'}\,ds\right]^{1/q'}\left[\int_R^\infty [K(s)]^q\,ds\right]^{1/q} \\ &\leq C\left[\int_R^\infty [K(s)]^q\,ds\right]^{1/q}, \end{aligned}$$

and then proceed as in the previous example. In particular, if $K(\omega) = O(\omega^{-m})$ with $m > 1$, and $f(\omega) = 1/(1 + |\omega|^p)$, $0 < p \leq 1$, choose $q' = \mu/p$, where $\mu > 1$ is to be determined. Then $f(\omega)$ is in the class $L^{q'}$, and we have only to verify that $(\omega^{-qm+1})^{1/q}\omega^p \to 0$ as $\omega \to \infty$, *i.e.*, that the exponent $-m + (1/q) + p = p + 1 - (p/\mu) - m$ is negative, which is clearly true when μ is close enough to 1. Hence, combining results, we find that example 1 is valid for $0 \leq p < 2$.

Asymptotic Behavior of $\Gamma(\omega,\omega')$

So far we have been concerned with the power spectrum $\gamma(\omega) = E\,|Y(\omega)|^2$. We now turn our attention to the *covariance* $\Gamma(\omega, \omega') = EY(\omega)\overline{Y}(\omega')$, given by

$$\Gamma(\omega, \omega') = \int M(\omega - s)\overline{M}(\omega' - s)f(s)\, ds. \qquad (4)$$

The asymptotic behavior of $\Gamma(\omega, \omega')$ is described by the following generalization of Theorem 1.

Theorem 2: Let $M(\omega)$ be square-integrable with autoconvolution

$$\mathfrak{M}(\omega) = \int_{-\infty}^{\infty} M(\omega + s)\overline{M}(s)\, ds, \qquad \mathfrak{M}(0) = 1.$$

Let $f(\omega)$ be a non-negative integrable function which is slowly varying and nonincreasing for sufficiently large ω. Suppose, as in Theorem 1, that

$$\lim_{\omega\to\infty} \frac{K(\omega)}{f(\omega)} = 0, \qquad (6)$$

where $K(\omega) \equiv |M(\omega)|^2$, and, in addition, suppose that $\mathfrak{M}(\omega)$ has a positive lower bound for $0 \leq \omega \leq A$. Then, if $\omega \leq \omega'' \leq \omega'$ all $\to \infty$ in such a way that $|\omega - \omega'| \leq A$, we have

$$\frac{\Gamma(\omega, \omega')}{f(\omega'')\mathfrak{M}(\omega - \omega')} \to 1 \qquad (20)$$

uniformly, where $\Gamma(\omega, \omega')$ is the covariance (4).

Remark 1: When $\omega = \omega'$, Theorem 2 becomes Theorem 1.

Remark 2: Both of the remarks made after Theorem 1 apply to Theorem 2 as well.

Remark 3: Of course Theorem 2 is susceptible to the same kind of extensions as Theorem 1.

Remark 4: If we choose $\omega'' = \frac{1}{2}(\omega + \omega')$, then (20) asserts that $\Gamma(\omega, \omega')$ is asymptotically the *locally stationary covariance* $f((\omega + \omega')/2)\ \mathfrak{M}\ (\omega - \omega')$.[12]

Proof of Theorem 2: Since the proof is almost identical to that of Theorem 1, we may be quite brief. We have

$$\Gamma(\omega, \omega') = \int_{-\infty}^{R} M(s)\overline{M}(s + \omega' - \omega)f(\omega - s)\, ds + \int_{R}^{\infty} M(s)\overline{M}(s + \omega' - \omega)f(\omega - s)\, ds = \Gamma_1(\omega, \omega') + \Gamma_2(\omega, \omega'),$$

where $R = R(\omega)$ is a function chosen as in Theorem 1. Applying the dominated convergence theorem to the first integral, we get

$$\lim_{\omega\to\infty} \frac{\Gamma_1(\omega, \omega')}{f(\omega)} = \int_{-\infty}^{\infty} M(s)\overline{M}(s + \omega' - \omega)\, ds = \mathfrak{M}(\omega - \omega'),$$

where in the left side $f(\omega)$ can be replaced by $f(\omega'')$ since $f(\omega)/f(\omega'') \to 1$. Then, applying the elementary inequality $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$ to the second integral, we have

$$|\Gamma_2(\omega, \omega')| \leq \frac{1}{2}\int_{R}^{\infty} \{|M(s)|^2 + |M(s + \omega' - \omega)|^2\} f(\omega - s)\, ds.$$

It follows that $\Gamma_2(\omega, \omega')/f(\omega)$, and hence $\Gamma_2(\omega, \omega')/f(\omega'')$ tends to zero precisely as in Theorem 1, where now $\frac{1}{2}\{|M(s)|^2 + |M(s + \omega' - \omega)|^2\}$ plays the role of $K(\omega)$.

Two Weighted Processes

There is another direction in which we can generalize Theorem 1, namely, we can consider the case of *two* weighted random processes. Thus, let $m_1(t)$ and $m_2(t)$ be two square-integrable weight functions and consider the two processes $y_1(t) = m_1(t)x(t)$ and $y_2(t) = m_2(t)x(t)$, both derived from the same underlying stationary process $x(t)$. Then, if $Y_1(\omega)$ and $Y_2(\omega)$ are the Fourier transforms of $y_1(t)$ and $y_2(t)$, which by a previous argument exist with probability one, we have

$$\gamma_{12}(\omega) = EY_1(\omega)\overline{Y}_2(\omega) = \int_{-\infty}^{\infty} M_1(\omega - s)\overline{M}_2(\omega - s)f(s)\, ds, \qquad (21)$$

where $M_1(\omega)$ and $M_2(\omega)$ are the Fourier transforms of $m_1(t)$ and $m_2(t)$. These considerations suggest the following theorem:

Theorem 3: Let $f(\omega)$ be a non-negative integrable function which is slowly varying and nonincreasing for sufficiently large ω. Let $M_1(\omega)$ and $M_2(\omega)$ be square-integrable functions such that

$$\lim_{\omega\to\infty} \frac{M_1(\omega)\overline{M}_2(\omega)}{f(\omega)} = 0.$$

Then

$$\lim_{\omega\to\infty} \frac{\gamma_{12}(\omega)}{f(\omega)} = \int_{-\infty}^{\infty} M_1(\omega)\overline{M}_2(\omega)\, d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} m_1(t)\bar{m}_2(t)\, dt,$$

where $\gamma_{12}(\omega)$ is the quantity (21).

Like Theorem 2, the proof of this theorem is essentially the same as that of Theorem 1, where this time we identify $K(\omega)$ with $M_1(\omega)\overline{M}_2(\omega)$. Again the two remarks made after Theorem 1 apply to Theorem 3 as well; moreover, Theorem 3 is susceptible to the same extensions as Theorem 1. The meaning of Theorem 3 is illustrated by the case where $m_2(t)$ is a translate of $m_1(t)$, *i.e.*, $m_2(t) = m_1(t + \tau)$. We take $m_1(t)$ to be a real-valued data window which peaks about some central value. Suppose that the two data windows overlap only slightly, *i.e.*, suppose that τ is such that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} m_1(t)m_1(t + \tau)\, dt \ll \frac{1}{2\pi}\int_{-\infty}^{\infty} m_1^2(t)\, dt = 1.$$

Then Theorem 3 asserts that $\lim_{\omega\to\infty}\{\gamma_{12}(\omega)/\gamma(\omega)\} \ll 1$, where as usual $\gamma(\omega)$ is the convolution (5). Qualitatively, this means that if we take two essentially non-overlapping sections of a stationary process $x(t)$, we shall find that their high-frequency contents are only weakly correlated, even if the two sections both lie *well within* the correlation distance or memory of $x(t)$, provided, of course, that appropriate conditions are met by $x(t)$ and the data windows involved.

[12] R. A. Silverman, "Locally stationary random processes," IRE Trans. on Information Theory, vol. IT-3, pp. 182–187; September, 1957.

Extremal Coding for Speech Transmission*

MAX V. MATHEWS†

Summary—A digital coding and its application to speech transmission is described. The coder determines the amplitudes and times of sucessive extremes (relative maxima and minima) of the signal. This information is decoded at the receiver by interpolating a function between extremes so as to connect them smoothly and preserve the extremes of the original signal in the reconstructed wave. Thus, the coding is a nonlinear sampling technique. It is related to clipped speech encoding which effectively transmits only the times of the extremes.

The properties of the coding for speech signals have been studied by digital simulation on an IBM 704 computer. Information rate, statistics of the extremes data, and quality of the resulting signal have been evaluated. The buffer size necessary to receive the randomly occurring data and transmit at a constant rate was measured.

Introduction

EXTREMAL coding is a simple procedure for representing a signal by a sequence of numbers. The coding can be considered for signals which a) are continuous functions of time, b) are to be transmitted digitally, and c) can tolerate a certain minimum distortion.

The transmitted information consists of the amplitudes of the signal at its extremes (a_1, a_2, $\cdots$ in Fig. 1) and the time intervals, $t_{i+1} - t_i$, between extremes. In the decoding process, a suitable function is interpolated between extremes as illustrated by the dashed curve in Fig. 1. The extremes of the original signal are preserved in the reconstructed wave, and there are no discontinuities in the wave or its first derivative at the extremes.

The coding is basically a nonlinear, signal dependent, sampling procedure. As such, it is not easily analyzed with presently available mathematics and is perhaps best evaluated by considering an application to a specific signal such as speech. The principal questions to be investigated are:

1) What is the information rate of the representation?
2) What is the effect of the inherent distortion?
3) What time delay is associated with the coder?

The time delay arises in a buffer which must be inserted between the randomly occurring extremal information source and any constant rate transmission facility.

This paper considers in detail the coding applied to speech signals. Application to television signals is presented elsewhere.[1] For speech, the coding is related to infinite clipping[2,3] in which the only transmitted information is the times of the extremes, these being the zero crossing times of the infinitely clipped derivative of the speech wave. In addition to these times, the extremal coding also sends the amplitudes of the wave at the extremes. Thus, better quality at the cost of a higher information rate is achieved.

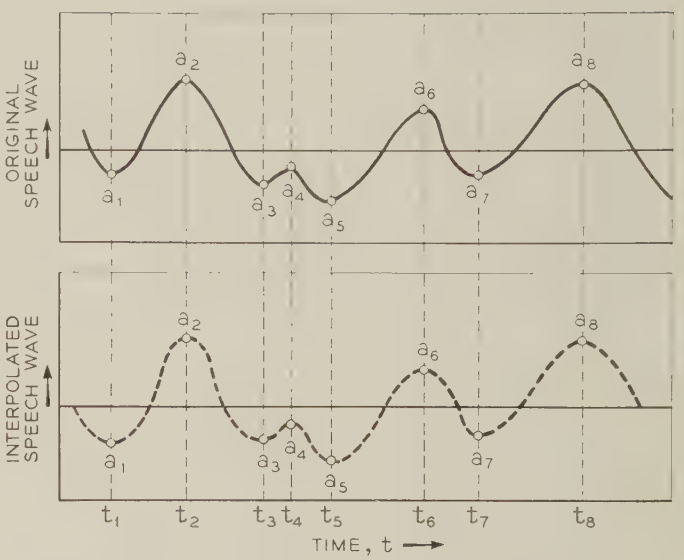


Fig. 1—Analysis and synthesis of extremal speech.

The study was carried out by digital simulation on an IBM 704 computer using speech-to-computer input-output equipment.[4] A number of speech samples for subjective evaluation were generated, and statistics of the extremes sufficient for channel capacity and buffer size estimates were measured.

Digital Simulation of Extremal Speech

The simulation program consisted of recording a digitalized speech passage on an IBM 704 magnetic tape and processing this tape with computer programs to obtain a number of digitalized samples representing the reconstructed extremal speech for various parameter values. In the process, the extremal channel capacity and buffer size were evaluated. The reconstructed samples were converted to audible speech for subjective evaluation.

A. Preparing Source Tape

The original speech passage contains four different sentences spoken simultaneously into a good quality dynamic microphone and a carbon button microphone

* Manuscript received by the PGIT, May 11, 1959.
† Bell Telephone Labs., Inc., Murray Hill, N. J.

[1] B. Julesz, "Coding television signals based on edge detection," *Bell Sys. Tech. J.*, vol. 38, pp. 1001–1020; July, 1959.

[2] J. C. R. Licklider and Irwin Pollack, "Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech," *J. Acoust. Soc. Am.*, vol. 20; January, 1948.

[3] J. C. R. Licklider, "Intelligibility of amplitude-dichotomized, time quantized speech waves," *J. Acoust. Soc. Am.*, vol. 22, pp. 820–823; November, 1950.

[4] E. E. David, Jr., M. V. Mathews, and H. S. McDonald, "Description and results of experiments with speech using digital computer simulation," IRE Wescon Convention Record, pt. 7; 1958.

(telephone handset). Each sentence was uttered by a different speaker, two being male and two being female. The duration of the total passage is 10.4 seconds. As a result of a digital recording error, one of the handset sentences was ruined so the final passage comprises four microphone sentences (10.4 seconds) and three handset sentences (8.9 seconds).

The speech was filtered with a 4-kc low-pass filter, sampled at 10,000 times per second, uniformly quantized into 1024 levels (10 bits) per sample and recorded on a digital computer tape which served as the input to all programs. Negative speech peaks were at level 0, positive peaks at level 1024 and *speech zero at level* 512.

B. Simulation Program

The simulation program contained two major parts, the first for locating the extremes, the second for interpolating a wave between the extremes. Each extreme was located by:

1) Selecting a sample which is an extreme with respect to the two adjacent samples.
2) Passing a quadratic polynomial through the extreme sample and the two adjacent samples.
3) Taking the peak of the polynomial as the extreme of the speech wave.

By this interpolation process, the extreme times t_i were determined to an accuracy greater than one sampling time. This fairly elaborate procedure for locating the extremes was found necessary to remove a quaver which occurred in the reconstructed speech if the extremes were located only to an accuracy of one sampling time.

The extreme amplitude a_i was further quantized into 32, 64, or 128 levels using nonuniform quantizing characteristics as illustrated in Fig. 2. The quantizing characteristics are those suggested by Smith[5] for PCM speech. The quantizing interval d_i varies approximately according to the relation

$$d_i \approx 1 + \mu \frac{P_i}{P_N} \qquad i = 1 \cdots N \tag{1}$$

where d_i and the abscissa points P_i are defined in Fig. 2, $2N$ is the number of levels, and μ is the compression ratio. Compression ratios of 30 db for the 128, 64, and 32 level quantizing, and 35 db for the 8 and 16 level quantizing were used. (The 8 and 16 level quantizing was used for PCM comparison which will be described later.)

The reconstructed speech was generated according to the relation

$$S'(t) = a_i + (a_{i+1} - a_i)F_j\left(\frac{t - t_i}{t_{i+1} - t_i}\right) \qquad t_i \leq t \leq t_{i+1} \tag{2}$$

in which $S'(t)$ is the reconstructed speech and F_j is the interpolation function. Two interpolation functions have been examined,

$$F_1(X) = \left(\sin \frac{\pi X}{2}\right)^2 \tag{3}$$

and

$$F_2(X) = X^2(3 - 2X). \tag{4}$$

Both functions have the properties that

$$F(0) = F'(0) = F'(1) = 0 \tag{5}$$

and

$$F(1) = 1. \tag{6}$$

Thus, both

a) preserve the extremes of the original speech in the reconstructed speech,
b) make the reconstructed speech and its derivative continuous at the extremes, and
c) introduce discontinuities in the second derivate at the extremes.

The two functions are very similar and produce speech which sounds the same. All the data reported in this paper were obtained with the $F_2(X)$ function.

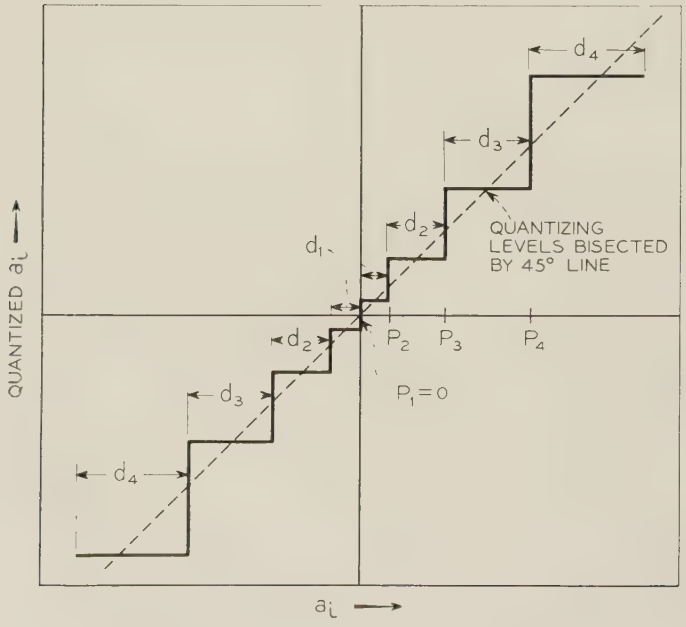


Fig. 2—Nonuniform quantizing characteristic.

C. Threshold and Bump Removal Parameters

The basic coding described above is modified by two parameters, one of which throws out extremes resulting from noise when no speech is present, the other of which eliminates speech wave bumps which are small compared with the surrounding topography. These modifications substantially reduce the number of extremes while only slightly affecting the quality of the speech.

[5] B. Smith, "Instantaneous compounding of quantized signals," *Bell Sys. Tech. J.*, vol. 36, pp. 653–709; May, 1957.

During the silent portions of speech, it is desirable to move the extremes resulting from background noise as these are frequent and represent no information. A program for this purpose examines groups of 10 consecutive extremes and if

$$\frac{1}{10}\sum_{i=10k}^{10k+9} | a_{i+1} - a_i | < T \qquad k = 1, 2, \cdots \tag{7}$$

where T is the threshold, for three consecutive groups (three consecutive values of k), the extremes in the center group are removed. The requirement that three consecutive groups be below threshold is included to preserve the beginnings and ends of low energy sounds which are near the noise level.

The selection of a threshold is a compromise between clipping low level sounds and reducing the number of extremes, a compromise which depends strongly on the signal-to-noise ratio of the speech. The selections used will be discussed when the data are discussed.

In addition to low level extremes, certain bumps may be removed. For example, if the bump represented by a_3 in Fig. 3 is small enough compared with the previous

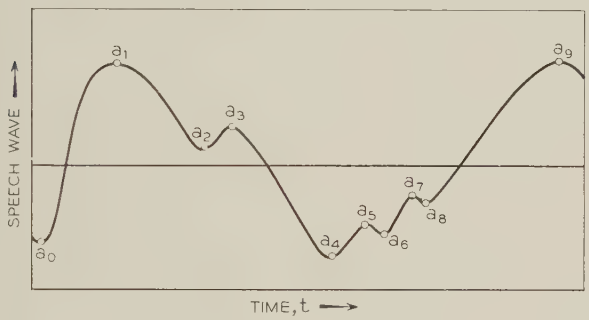


Fig. 3—Plot to illustrate bump removal.

fill a_0a_1, then a_2 and a_3 may be eliminated with little effect on speech quality. The essence of the bump remover program is to test

$$| a_{i+2} - a_{i+1} | < B | a_{i+1} - a_i | \tag{8}$$

where B is the bump fraction ($0 < B < 1$) and to remove the a_{i+1} and a_{i+2} extremes if the inequality is satisfied. The actual program has additional features which will not be described in detail, so that in a sequence of subthreshold extremes ($a_4 - a_8$ in Fig. 3), the minimum or maximum extreme (a_4) will be preserved and the others erased ($a_5 - a_8$). Also, in an area of rapidly decreasing amplitude, large hills are not allowed to erase subsequent bumps which are small only because they form the beginning of a low amplitude section.

D. Distribution of Extremes and Channel Capacity

During the simulation, the average rate of extremes, the distribution of interval $t_{i+1} - t_i$, and the entropy of this distribution were computed. The distributions (unnormalized) for most of the values of threshold T and bump fraction B which were computed are shown in Fig. 4. Additional distributions of somewhat different speech functions have been given by Davenport.[6]

The distributions for the microphone samples are similar, all having a principal peak at 0.2 –0.3 msec and a subsidiary but significant peak at 1.0–1.1 msec. The handset distributions are also similar among themselves and have their main peak at 0.2–0.3 msec. Comparing directly the microphone and handset distributions, the latter have a stronger component at 0.1–0.2 msec. As a consequence, the average rate of handset extremes is substantially higher than that of the microphone extremes. The source of this effect is not yet known, and the only conclusion that can be drawn now is that carbon buttons in their present state are not desirable for extremal coders.

From the distributions, an upper limit, H, on the average information content of the quantized $t_{i+1} - t_i$ intervals was estimated according to the relation

$$H = \left[\sum_{i=1}^{65} \frac{N_i}{N_T} \log_2 \frac{N_T}{N_i}\right] + 3 \tag{9}$$

where N_i is the number of intervals whose duration lies between $i - 1$ and i sample times, and N_T is the total number of intervals. The estimate neglects any dependence between successive intervals. The additional 3 bits per interval is included in the relation because the interval is quantized to 1/8 sampling time.

Some data on the joint distribution of $a_{i+1} - a_i$, $t_{i+1} - t_i$ which is not shown were also taken. The sample indicates that correlation between these two quantities is small, and little advantage can be taken of the correlation for coding.

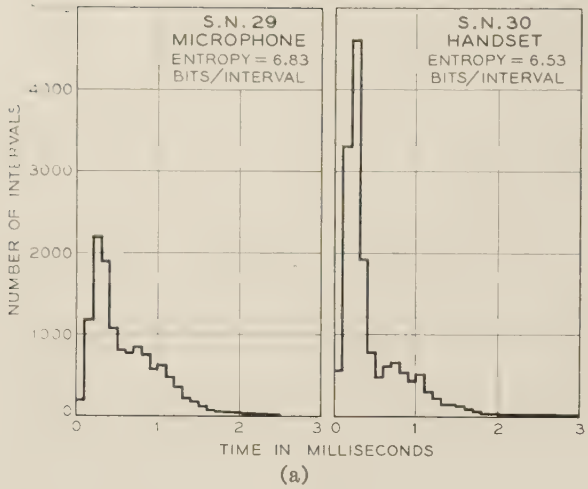
An upper limit on the channel capacity for the extremal data was computed by the equation,

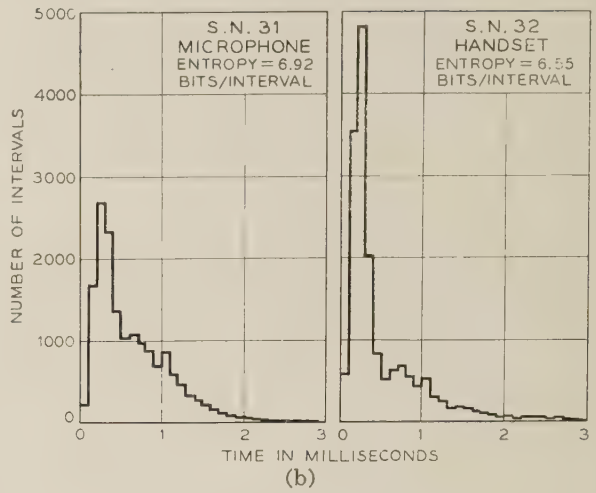
$$\text{channel capacity} = (\text{Average Rate of Extremes}) \cdot (\log_2 Q + H) \tag{10}$$

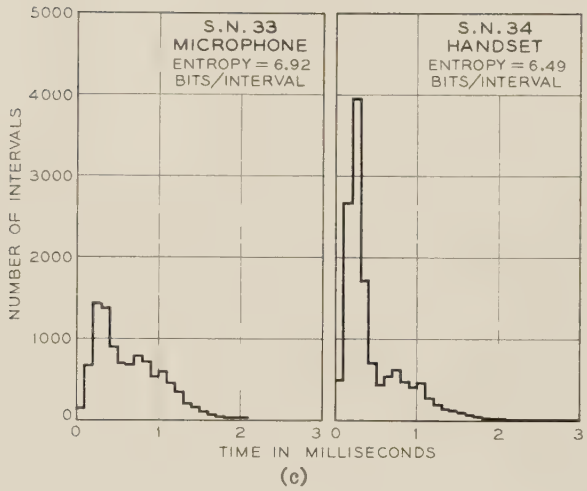
where Q is the number of quantizing levels of a_i and H is the average information content of the time interval distribution. The channel capacity estimate assumes that while additional coding will be done on the time interval data to take advantage of their distribution function, no further coding of the a_i's will be done. This rather arbitrary assumption was justified because the nonuniformly quantized a_i's have a nearly flat distribution.

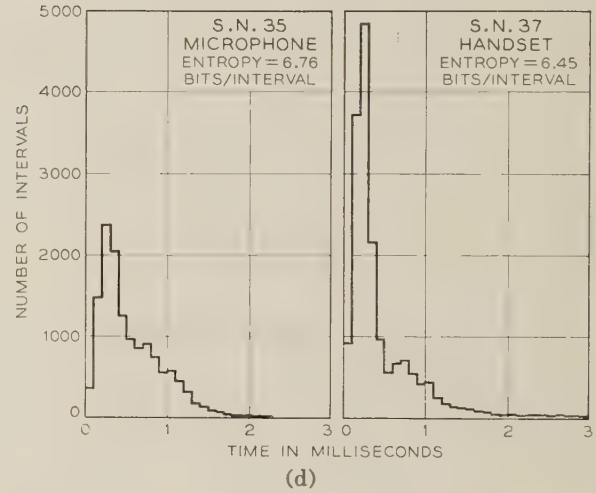
The computed data, samples 25 through 38, are summarized in Table I where the parameters of the

[6] W. B. Davenport, Jr., 'Experimental study of speech-wave probability distributions," *J. Acoust. Soc. Am.*, vol. 24, pp. 390–399; July, 1952.









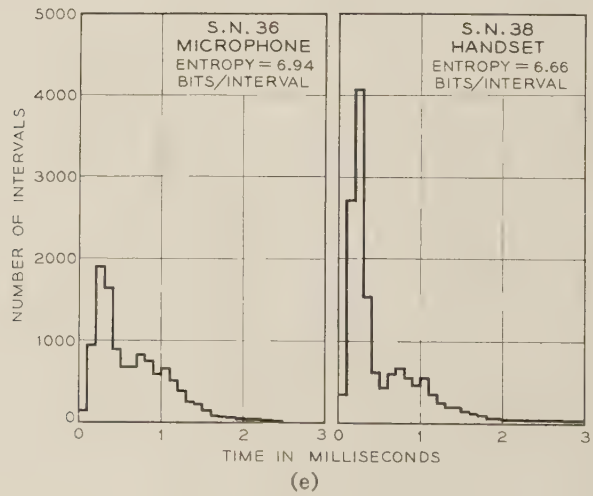


Fig. 4—Distribution of intervals between extremes. (*a*) $Q = 128$, $T = 4.5$, $B = 3/32$, (*b*) $Q = 128$, $T = 0$, $B = 3/32$, (*c*) $Q = 128$, $T = 9.0$, $B = 3/32$, (*d*) $Q = 128$, $T = 4.5$, $B = 0$, (*e*) $Q = 128$, $T = 4.5$, $B = 3/16$.

TABLE I
SUMMARY OF EXTREMAL DATA

Sample No.	Source Type	Quantizing levels	Threshold T (Maximum speech signal = 500)	Bump Fraction B	Inf. Content of time interval Bits per sample	Extremes per second	Channel capacity Bits per sec.
25	Microphone	32	4.5	3/32	6.84	1225	14500
26	Handset	32	4.5	3/32	6.58	1774	20500
27	Microphone	64	5.4	3/32	6.84	1215	15600
28	Handset	64	4.5	3/32	6.53	1816	22700
29	Microphone	128	4.5	3/32	6.83	1220	16900
30	Handset	128	4.5	3/32	6.53	1830	24700
31	Microphone	128	0.0	3/32	6.92	1572	21900
32	Handset	128	0.0	3/32	6.55	1918	26000
33	Microphone	128	9.0	3/32	6.92	983	13690
34	Handset	128	9.0	3/32	6.49	1575	21200
35	Microphone	128	4.5	0	6.76	1318	18130
36	Microphone	128	4.5	3/16	6.94	1125	15700
37	Handset	128	4.5	0	6.45	1984	26700
38	Handset	128	4.5	3/16	6.66	1647	22500
41	Microphone	64	4.5	3/16	6.94	1128	14600
42	Handset	64	4.5	3/16	6.67	1650	20900
44	Microphone	128	2.0	1/16	6.84	1480	20500
44	Handset	128	4.0	1/16	6.52	1901	25700

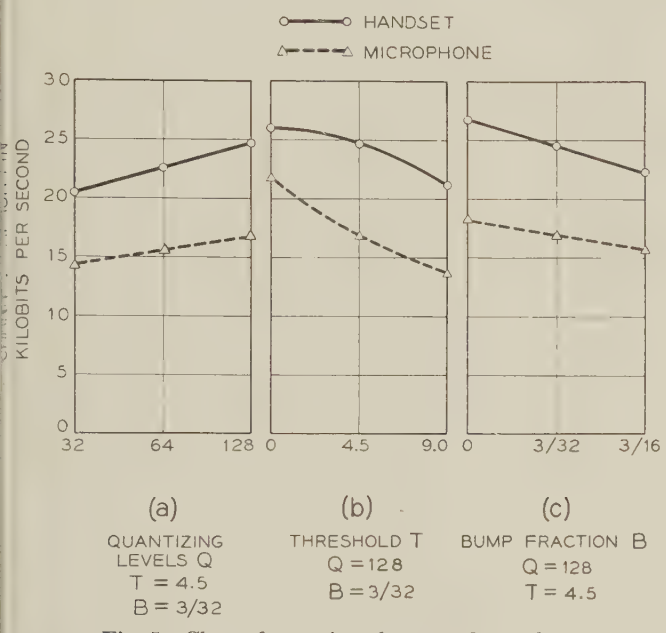


Fig. 5—Channel capacity of extremal speech.

arious conditions studied are listed along with the esulting average information content of extreme interval, verage rate of extremes, and channel capacity. (Also sted are samples 41 through 44, which will be discussed ater.) The channel capacity is plotted in Fig. 5.

Prior to the subjective tests described in the next ection, certain qualitative judgments were formed by stening to the samples critically. Most of the listeners ad considerable previous listening experience and were amiliar with the general format of the coding scheme.

1) As a function of quantizing levels, the quality improves substantially going from 32 to 64 levels but not much from 64 to 128 levels. Since the channel capacity increases uniformly with the logarithm of number of quantizing levels, 64 levels appears to be the most efficient selection.
2) The low level clipping became objectionable with a T of 4.5 in the microphone case and with a T of 9.0 in the handset case. The absolute levels of T depend so strongly on the background noise that they are not by themselves significant. A more meaningful figure is the reduction in channel capacity before objectionable clipping. The reduction is about 20 per cent in both cases.
3) The bump threshold degradation was detectable only for B of 3/16; thus 15 per cent of the extremes representing minor bumps can be removed with almost no degradation. Such a removal is a substantial modification of the waveform and the small amount of degradation so introduced is surprising.

On the basis of these observations, four additional samples (41–44) were generated for more exact subjective evaluation. The first two samples were selected to have a low channel capacity with reasonable quality by setting $Q = 64$, T, 4.5 = and $B = 3/16$. The second pair were selected as a high channel capacity, high quality case with $Q = 64$, $B = 1/16$, and $T = 2.0$ (Microphone), $T = 4.0$ (Telephone). The subjective comparisons with PCM speech are described below.

E. Buffer Size

The size of buffer required to receive the randomly occurring extremes and transmit extremal information at a constant rate is a complicated function of the statistics of the extremes. The function depends on such high order statistics that it was believed a better estimate of size could be obtained from a direct measurement for the samples at hand rather than from a statistical estimation. Consequently, an approximate measurement was made during the simulation.

The number of samples $B(t)$ in the buffer at any time may be written

$$B(t) = E(t) - rt$$

where $E(t)$ is total number of extremes occurring between times 0 and t as illustrated in Fig. 6, and r is the rate of taking extremes from the buffer.[7] The rate r must be close to the average rate of extremes

$$r = \frac{E(T)}{T}$$

where T is the time over which the buffer is to be evaluated. The size of the buffer will then be the maximum minus the minimum contents:

$$\text{Buffer Size} = B(t)_{\text{Max}} - B(t)_{\text{Min}}.$$

This function was evaluated by the computer.

Plots of buffer size as a function of low level threshold T and bump fraction B are given in Fig. 7. The size appears to be a random function of both these parameters and has no discernable trend. Such might be expected since the size is the difference of the absolute maxima and absolute minima of the difference between two functions, one of which fluctuates rapidly, as illustrated in Fig. 6. Therefore, a small change in $E(t)$ may have great and unpredictable effects on the buffer size.

The average size is about 1400 extremes. At an approximate rate of 1400 extremes per second, the buffer will introduce a delay in the order of one second into the transmitted speech. Such a delay is several times the maximum tolerable delay in two-way communication, and would preclude such a use for the coding. The delay, however, could be greatly reduced in a multichannel system where the extremes of many speakers would combine into a more nearly uniform flow.

Subjective Comparison with PCM

A subjective comparison of the extremal speech with companded PCM speech was made to estimate a quality equivalence between these codings and, thus, to estimate the efficiency of the extremal coding relative to PCM. The comparison was designed to elicit a preference statement as a measure of quality.

The test was carried out by dubbing 3 PCM samples and one extremal sample on 4 tracks of a 14-track tape recorder. Each sample contains the 4 previously described sentences (3 sentences in the case of telephone speech). The PCM samples were quantized respectively into 8, 16 and 32 levels and sampled 10,000 times per second giving pulse rates of 30,000, 40,000 and 50,000 per second. The *same digital source* tape was used for the PCM and extremal codings so that *all* differences between samples resulted from the coding. The audio tape was formed into a loop so the samples repeat, the starting points of the samples being synchronized in all channels.

[7] With this definition, $B(t)$ may be negative, indicating a negative number of samples in the buffer. This situation is of course impossible. The actual buffer will contain ($B(t)$ + Constant) samples. However, for this computation, the constant can be neglected as it does not affect the estimate of the buffer size.

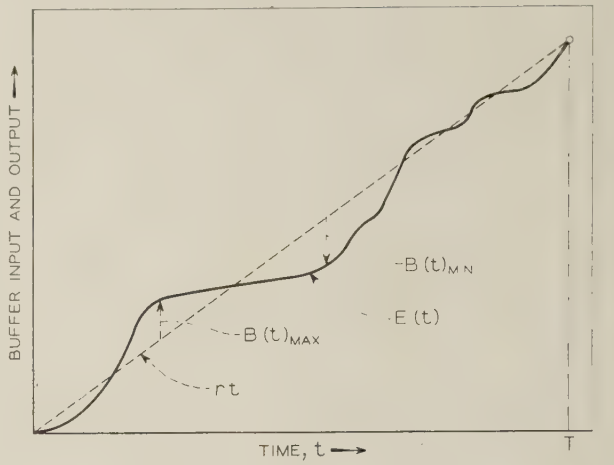


Fig. 6—Plot showing buffer contents.

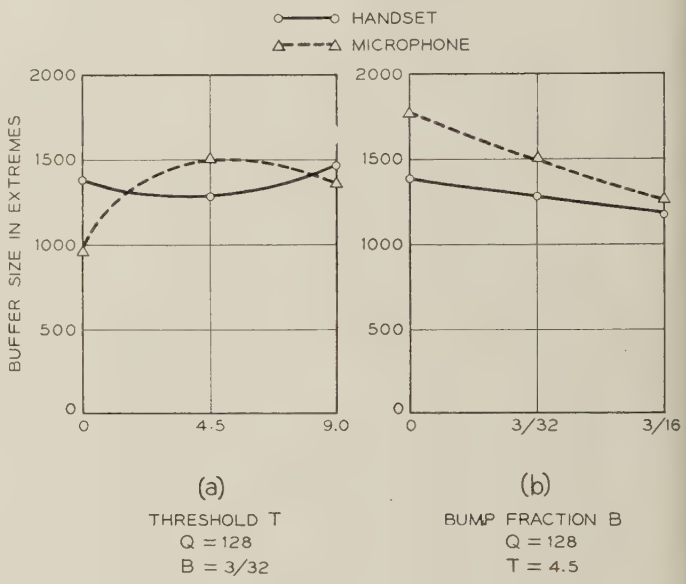


Fig. 7—Buffer size.

In a test, the samples were presented to the subject in pairs, using the apparatus shown in Fig. 8. The experimenter controlled which samples comprised a given pair, and the subject could listen to either member by means of a button which switched the output each time it was pushed. The subject was encouraged to listen as long as necessary, alternating channels as often as desired until he could answer the question: "Which of the two samples do you like best or, in other words, which would you prefer to listen to on the telephone?" The samples were then changed until all possible pairs of the 4 signals had been presented twice to the subject (a total of 12 responses). In each response, the preferred channel was ranked 1, the other 2. Five or six subjects were run for each test, the subjects being girls who were technically naive and completely unfamiliar with the coding. The final test results presented in Table II are the summation of the rankings for each coding.

Four tests were run, two with microphone, two with handset speech. In the results, a low ranking indicates a

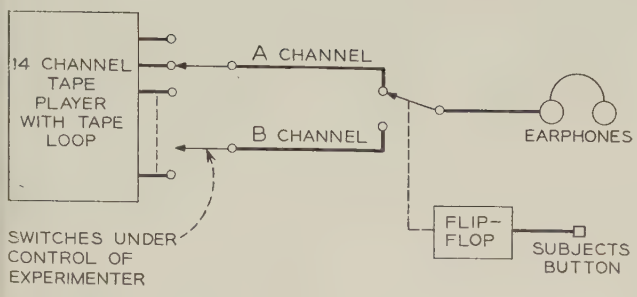


Fig. 8—Subjective test setup.

An Evaluation of Extremal Coding

The conclusions which may be drawn from this simulation can be summarized as follows:

1) Extremal codings can be considered for applications requiring digital transmission of speech.
2) In comparison with companded PCM, the extremal coding requires about half the channel capacity of PCM for equivalent quality transmission. This factor is the maximum achieved over a range of pulse rates and with a dynamic microphone source.

TABLE II
Subjective Test Results

Extremal Sample No.	Number of Subjects	Speech Source	Channel capacity Bits/sec.	Ranking				
				Extremal	30,000 Bit/sec PCM*	40,000 Bit/sec PCM*	50,000 Bit/sec PCM*	Confidence Rating
41	6	Microphone	14600	59	62	49	46	.040
43	6	Microphone	20500	54	68	53	41	.004
42	5	Handset	20900	54	49	40	37	.011
44	5	Handset	25700	45	57	38	40	.004

* 10,000 samples per second

preferred sample. Both the low (no. 41) and high (no. 43) bit rate microphone samples ranked between the 30,000 and 40,000 bit/sec PCM, the low being closest to the 30,000 PCM and the high being closest to the 40,000 PCM. Thus, a channel capacity reduction of about 2.2 over PCM is achieved by the low rate extremal and only about 1.9 by the high rate extremal. On the basis of this result, it would seem that the inherent extremal degradation is such that highest efficiency occurs for low quality transmission at low bit rates.

The handset samples, (nos. 42 and 44) by virtue of their higher extreme rates, have markedly lower efficiencies than the microphone samples, and lower quality compared with PCM. Thus, the carbon button should definitely be avoided as an extremal source.

The statistical significance of the paired comparisons have been evaluated.[8] With the principal assumptions that the ranking is carried out on a one dimensional psychological scale and that successive rankings are independent, the probability that any of the observed spreads of ranking would occur by chance is less than 0.04. Consequently, the data seems sufficiently reliable. In addition, in all except the test of sample 44, the order of preference among PCM samples is also the order of increasing channel capacity. This ordering lends additional credence to the results.

The subjective comparisons indicate that a factor of about two can be saved by extremal coding for speech transmission, which is comparable to 30,000 bits per second PCM.

3) The inherent degradation in the extreme interpolation function is such that extremal coding is efficient only at low channel capacities. Hence, little improvement in quality is achieved by increasing the rate above 15,000 bits per second. This conclusion is reinforced by two simulations not reported in detail here. The first applied extremal coding to differentiated speech. The second used a more extensive interpolation rule which removed the discontinuities from the second derivative at the extremes. These modifications did not substantially improve the speech quality.
4) In comparison with infinitely clipped speech, the extremal coding produces a better quality transmission using about three times the channel capacity (the amplitude information being additional in the extremal coding). The data obtained from the extremal simulation can also be used to obtain a better estimate of the information rate of the infinitely clipped speech. Thus, assuming a 100 μsec time quantization, the $t_{i+1} - t_i$ distributions (Fig. 4) yield information rate estimates ranging from 4700 bits per second to 6170 bits per second.
5) A substantial buffer of about 1500 extremes or 15,000 bits and associated transmission time delay of about one second is required for a single channel coder. This requirement contra-indicates the use of single channel extremal coders. However, by taking advantage of the more uniform statistics of a multichannel system, by not requiring a completely uniform flow of information, and by tolerating a certain percentage of saturation, both the buffer and the time delay can be eliminated.

[8] R. A. Bradley and M. E. Terry, "Rank analysis of incomplete block designs," *Biometrika*, vol. 39, pts. 3 and 4; December, 1952.

The extremal coding is an example of a class of processes which represent the speech time waveform as a sequence of simple features. Such codings achieve reductions in channel capacity because either a) the features occur less frequently than the Nyquist rate or b) a set of features are indistinguishable to the ear. The limit of channel capacity reduction achieved by the direct codings already tried is 3 to 4 and it is doubtful whether this limit will be much increased by future work.

One side point should be mentioned in this summary. The extremal study was facilitated by digital simulation. Without this tool it would not have been possible to carry out the detailed comparisons with the precision described here. The extremely important questions of information rate and buffer size could not have been approached so directly, and of course, the cost and toil involved in constructing an extensive laboratory model were avoided. Hence, it seems unlikely that the extremal study would have been considered feasible without simulation, and simulation demonstrated its effectiveness in this study.

Correspondence

On Periodicity of States in Linear Modular Sequential Circuits*

A number of investigators have been concerned with sequential circuits comprising ideal delays, modulo p summers, and amplifiers with gain equal to integers $-<p=$ prime. It is convenient to call a device of this type a *linear modular sequential circuit* (LMSC) since it may be characterized by the linear relations [1]

$$\mathbf{y}(n) = C\mathbf{s}(n) + D\mathbf{x}(n),$$
$$\mathbf{s}(n+1) = A\mathbf{s}(n) + B\mathbf{x}(n), \qquad (1)$$

where $\mathbf{x}(n)$, $\mathbf{y}(n)$, $\mathbf{s}(n)$ are input, output and state vectors, and A, B, C, and D are matrices; all are defined on the *modular field* $GF(p)$; (*i.e.*, the integers 0, 1, $\cdots$, $p-1$ and the operations $+$ and $\times$ modulo p).

The linear binary sequential circuits ($p=2$) were extensively investigated by Zierler [2] and Huffman [3]. The latter also briefly considered ternary circuits. Hartmanis [4] considered the general case. Huffman demonstrated that a pair of binary circuits may be used as coding and decoding filters for binary error correcting codes [5]. In extending this method of coding to multiple level codes, some questions concerning the periodicity of the states in an unexcited LMSC have arisen which have motivated this investigation. An unexcited LMSC is governed by the relation

$$\mathbf{s}(n+1) = A\mathbf{s}(n), \qquad (2)$$

where A is an $r \times r$ matrix ($r=$ *order* of the LMSC), called the *characteristic* matrix. The *state period* t is the smallest integer ν for which $s(n+\nu) = s(n)$. Since, from (2), $s(n+\nu) = A^{\nu}s(n)$, it is clear that the nullity of $A^t - I > 0$. If $t = T$ is such an integer that the nullity of $A^T - I$ is equal to r(*i.e.*, $A^T = I$), T is called the *matrix period*. The *minimum polynomial* $m(x)$ of A is the polynomial of lowest degree k so that $m(A) = 0$.

Lemma 1: (Fundamental Lemma) If $m(x)$ is irreducible, then the set of polynomials in A of degree $k-1$ over $GF(p)$ form a Galois field $GF(p^k)$ of order k.

Proof: Form the Galois field $GF(p^k) = GF[p, m(\lambda)]$; denote the elements of this field by

$$P_i(\lambda) = a_k^i + a_{k-1}^i + \cdots + a_1^i\lambda^{k-1} \quad a_j^i \in GF(p).$$

We exhibit the isomorphism $P_i(\lambda) \leftrightarrow P_i(A)$, where $P_i(A)$ is a matrix polynomial in A with coefficients corresponding to those of $P_i(\lambda)$. It is obvious that $P_u(A) + P_v(A) = P_w(A) \leftrightarrow P_w(\lambda)$. Consider $P_q(A)P_r(A)$. This product is a matric polynomial of degree $2(k-1)$ which may be reduced, using $m(A) = 0$, to a polynomial of degree $k-1$. The corresponding product $P_r(\lambda)P_q(\lambda)$ has the identical coefficients since λ is a root in $GF[p, m(\lambda)]$ of $m(\lambda)$. Hence

$$P_q(A)P_r(A) = P_s(A) \leftrightarrow P_s(\lambda) = P_r(\lambda)P_s(\lambda),$$

completing the proof.

Theorem 1: If $m(x)$ is irreducible, the matrix period of A is equal to the order T of the subgroup generated by λ in $GF[p, m(\lambda)]$, and every state period is equal to T. (The zero state is excluded).

Proof: Since $A \leftrightarrow \lambda$, $A^{\nu} \leftrightarrow \lambda^{\nu}$. Then, if T is the order of the subgroup generated by λ in $GF[p, m(\lambda)]$, $\lambda^T = 1$, hence $A^T = I$. Moreover, $A^{\nu} - I (\nu < T)$ may be reduced to a non-zero polynomial in A of degree $k-1$ which, being an element of a Galois field, is nonisingular. This establishes the theorem.

Corollary 1.1: The state and matrix period of a circuit whose minimum polynomial is irreducible and of degree k is a divisor of $p^k - 1$.

Corollary 1.2: There exists an LMSC for every modulus p and every order k so that the state and matrix period are of maximum length $p^k - 1$.

(The problem of finding a LMSC with maximum-length period entails the determination of a polynomial $m(\lambda)$ so that λ is a primitive member of $GF[p, m(\lambda)]$).

Example 1. Consider the LMSC with a characteristic matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} \bmod 3.$$

The minimum polynomial of A is $m(x) = x^2 + 1$ which is irreducible in $GF(3)$. Hence $A \leftrightarrow \lambda$ in $GF[3, \lambda^2 + 1]$. It can be shown that λ is 4th order in $GF[3, \lambda^2 + 1]$, hence $T = 4$.

If the minimum polynomial of A is reducible [over $GF(p)$], the set of polynomials in A of degree $k-1$ do not comprise a Galois field. It is not possible to assert that $A^{\tau} - I$ is either 0 or nonsingular; consequently, states lie in sequences of various lengths. It is clear, however, that every state period t is a divisor of the matrix period T, and hence if $T =$ prime, $t = 1$ or T. To investigate the periodicity of states in LMSC's which have reducible minimum polynomials we establish:

Lemma 2: If A is similar to $B =$ the direct sum of matrices B_i ($i = 1, 2, \cdots, s$) each with period R_i, then the period of A is $T =$ l.c.m. $(R_1, R_2, \cdots, R_s)$.

Proof: There exists a (nonsingular) matrix M so that $A = MBM^{-1}$ where $B =$ diag $[B_1, B_2, \cdots, B_s]$ (the direct sum). Now $A^T = MB^TM^{-1}$, hence, $A^T = I$ if and only if $B^T = I$. Now $B^{\nu} =$ diag $[B_1^{\nu}, B_2^{\nu}, \cdots, B_s^{\nu}]$. Thus $B^{\nu} = I$ if and only if $B_i^{\nu} = I$ ($i = 1, 2, \cdots, s$). The smallest integer ν for which this occurs is $\nu = T =$ l.c.m. $(R_1, R_2, \cdots, R_s)$, which establishes the lemma.

Theorem 2: If the factors $m_i(x)$ ($i = 1, 2, \cdots, s$) (over $GF(p)$) of the minimum

* Received by PGIT, October 20, 1958.

olynomial are distinct, and T_i = order of ubgroup generated by λ in $GF[p, m_i(\lambda)]$, hen the period of A is T = l.c.m. $(T_1, T_2, \cdots, T_s)$.

Proof: Under the conditions of the heorem, each matrix B_i of Lemma 2 may e chosen to satisfy $m_i(B_i) = 0$. From Theorem 1, the period of each B_i equals he order of the subgroup generated by in $GF[p, m_i(\lambda)]$. Hence, by Lemma 1, he theorem is established.

It can easily be shown that the di-ensionality of the subspace of states aving $t = T_i$ is equal to the rank of the natrices B_i in Theorem 2.

Example 2. Consider the LMSC whose haracteristic matrix is given by

$$A = \left[\begin{array}{cc|ccc} 1 & 1 & & 0 & \\ 1 & 0 & & & \\ \hline & & 0 & 1 & 1 \\ & 0 & 1 & 0 & 0 \\ & & 0 & 1 & 0 \end{array}\right] \bmod 2$$

Here $p = 2$ and $k = 5$ and there are $2^5 = 32$ tates in the state space. The characteristic nd minimal polynomial is

$$m(x) = (x^2 + x + 1) \cdot (x^3 + x + 1) \bmod 2$$

nd each factor is irreducible in $GF(2)$. We ust thus examine the period of the sub-roup generated by λ in both $GF[2, \lambda^2 + + 1]$ and $GF[2, \lambda^3 + \lambda + 1]$. It is found hat in each group λ is primitive; hence, he corresponding periods are $t_1 = 2^2 - = 3$ and $t_2 = 2^3 - 1 = 7$. Using Theorem , the matrix period is

$$T = \text{l.c.m. } [3, 7] = 21$$

nd every state period is a divisor of 21. n particular, there is a 2-dimensional ubspace of states (3 states) having a state eriod of 3, resulting in one cycle of length , and a 3-dimensional subspace of states 7 states) having a state period of 7, re-ulting in one cycle of length 7. The emaining 21 states lie in a single cycle f length 21.

The case in which there are repeated actors in the minimum polynomial of A s governed by:

Theorem 3: If the minimum polynomial $m(x) = \Pi_{i=1}^{q} m_i(x)^{e_i}$, and T_i = order of ubgroup generated by λ in $GF[p, m_i(\lambda)]$, hen

$$T = \text{l.c.m. } (k_1 T_1, k_2 T_2, \cdots, k_q T_q)$$

$_i$ = integer ≥ 1.

Proof: Each matrix B_i of Lemma 2 may e written in the rational canonical form [7]

$$B_i = \begin{bmatrix} C_i & U & \cdots & 0 & 0 \\ 0 & C_i & \cdots & 0 & 0 \\ & \cdots & & & \\ 0 & 0 & \cdots & C_i & U \\ 0 & 0 & \cdots & 0 & C_i \end{bmatrix}$$

where C_i is a matrix having $m_i(x)$ as its minimum polynomial; U is a matrix whose elements are all zero except the lower left hand element which is 1. (The number of rows of $B_i \leq e_i$.) Now

$$B_i^{\nu} = \begin{bmatrix} C_i^{\nu} & N_{12} & \cdots & N_{1e} \\ 0 & C_i^{\nu} & \cdots & N_{2e} \\ & \cdots & & \\ 0 & 0 & \cdots & C_i^{\nu} \end{bmatrix}$$

(where N_{ij} is a function of C_i and U). Therefore, a necessary condition for $B_i^{\nu} = I$ is $C_i^{\nu} = I$; hence, $R_i = k_i T_i$.

In summary, these results represent a fairly straightforward application of Galois field theory to a specific type of sequential circuit. The writers believe that other problems in discrete systems and coding may be treated by the use of the theory of finite fields.

Bernard Friedland
Thomas E. Stern
Dept. of Elec. Eng.
Columbia University
New York, N. Y.

References

[1] B. Friedland, "Linear modular sequential circuits," IRE Trans. on Circuit Theory, vol. CT-6, pp. 61–68; March, 1959.
[2] N. Zierler, "Several Binary-Sequence Generators," MIT Lincoln Lab. Tech. Rep. No. 95; September 12, 1955.
[3] D. A. Huffman, "The Synthesis of Linear Sequential Coding Networks," in "Information Theory," C. Cherry, ed., Academic Press, Inc., New York, N. Y., pp. 77–95; 1956.
[4] J. Hartmanis, "Linear Multivalued Sequential-Coding Networks," General Electric Res. Lab. Rep. No. 57-RL-1835, Schenectady, N. Y.; November, 1957.
5] D. A. Huffman, "A linear circuit viewpoint on error-correcting codes," IRE Trans. on Information Theory, vol. IT-2, pp. 20–28; September, 1956.
[6] T. E. Stern and B. Friedland, "Application of modular sequential circuits to single-error-correcting p-nary codes," this issue, pp. 114–123.
[7] S. Perlis, "Theory of Matrices," Addison-Wesley Press, Cambridge, Mass.; 1952.

Poincaré, Metric Reliability and Switching Components*

More than fifty years ago the French mathematician, Jules Henri Poincaré[1] contrasted what he called the mathematical continuum and the physical continuum. It was his impression that only the mathematical continuum had a transitive relation[2] of equality and that the physical continuum did not have such a transitive relation of equality.

This brief note serves to quantify Poincaré's notions of continuums in the context of probability theory, metric spaces, and the unifying branch of engineering mathematics called stochastic switching circuits.[3–5]

1) Consider the set α of states of some switching component (the cardinality[6] of α is not to exceed aleph-null). Then, α is said to be the state-space of the component. Identify corresponding states of excitation and response of the switching component by the same positive integer. If $A \in \alpha$ is a steady-state of excitation of the component, $B \in \alpha$ is the steady-state of response of the component resulting from the state A of excitation, and Λ_B is the set of response states of the component different from B, then let $P(A, b)$, $b \in \Lambda_B$ be the probability that b corresponds to A.

If $\sum_{b \in \Lambda_B} P(A, b)$ exists, call it $P_\epsilon(A, B)$. The function P_ϵ is said to be the component probability of error. We may also think of P_ϵ as the probability of being able to differentiate between A and B in the steady-state and with respect to the identification of excitation and response states given previously.

2) Based on this motivation we make the following postulate:

(a) $P_\epsilon(A, B) = 1$, if and only if $A = B$; for symmetry we postulate

(b) $P_\epsilon(A, B) = P_\epsilon(B, A)$, $A \in \alpha$, $B \in \alpha$; and wishing to reduce the effects of transitivity, we postulate

(c) $P_\epsilon(A, B)\, P_\epsilon(B, C) \leq P_\epsilon(A, C)$, $A \in \alpha$, $B \in \alpha$, $C \in \alpha$.

3) Put $r(A, B) = -\ln P_\epsilon(A, B)$ and call r the component reliability function. The intuitive significance of r is apparent. As $P_\epsilon \to 0$ then $r \to +\infty$ and as $P_\epsilon \to 1$ then $r \to 0$. In addition, r has the usual entropy connotations. From section 2 of this note, it is a simple matter to show that r is a Fréchet metric[7] on the state-space of the component and, hence,[7] that $_r = r/(1 + r)$ is a bounded metric on the state-space of the component.

4) Further mathematical significance can be given to the previous discussion. Thus, there is a context in which it is possible to discuss the reliability corresponding to a metrization of a curve in some space.

While working on the material of this note the author was supported, in part, by The United States Office of Naval Research.

A. A. Mullin
University of Illinois
Urbana, Ill.

* Received by the PGIT, January 8, 1959.

[1] J. H. Poincaré, "Science and Hypothesis," The Science Press, New York, N. Y., pp. 17–28; 1905.
[2] G. Birkhoff and S. MacLane, "A Survey of Modern Algebra," The Macmillan Co., New York, N. Y., p. 3; 1953.
[3] E. F. Moore and C. E. Shannon, "Reliable circuits using less reliable relays," *J. Franklin Inst.*, vol. 262, pt. I, pp. 191–208; September, 1956, and pt. 2, pp. 281–297; October, 1956.
[4] A. A. Mullin, "Reliable stochastic sequential switching circuits," *Trans. AIEE (Commun. and Electronics)*, vol. 77, pp. 606–611; November, 1958.
[5] A. A. Mullin, "Stochastic combinational relay switching circuits and reliability," IRE Trans. on Circuit Theory, vol. CT-6, pp. 131–133; March, 1959.
[6] F. Hausdorff, "Set Theory," Chelsea Press, New York, N. Y., pp. 28–29; 1957.
[7] J. L. Kelley, "General Topology," D. Van Nostrand Co., Inc., New York, N. Y., pp. 118–121; 1955.

Optimal Properties in the Statistical Theory of Reception*

Many authors have discussed various applications of the Bayes' inverse probability theorem to the problem of optimum receiver design. These fall roughly into two not completely mutually exclusive classes: those which might be classed as pure "detectors" and are expected to make decisions as to the presence or absence of a signal; and those which are expected to estimate the best value of a continuous parameter or to continuously demodulate a continuously modulated signal.

Siegert's "Ideal Observer"[1] for the detection of radar echoes is an example of the first. The "Ideal Observer" may be viewed as one which makes successive decisions, in favor of that possibility which corresponds to a maximum *a posteriori* probability; this decision method minimizes the average rate of making mistakes. Modifications of this type of detector have been extensively discussed by Middleton[2] and many others. In all these instances, the detector may be said to be optimum in some sense, provided the required statistical *a priori* signal information is correct. Woodward[3] and Davies, DuWaldt[4] and others have discussed the application of the inverse probability theorem to the radar location problem, although the former point out that, at least in the military radar problem, a systematic scheme for choosing a "best" estimate of position is rather hard to achieve due to the usual paucity of *a priori* knowledge concerning the possibility of a target. Their "best" output, if possible, is the complete *a posteriori* distribution function. Choosing the maximum of this distribution for a "best" estimate is an obvious possibility, but they give no justification, in terms of the quality of results obtained, for such a choice. Lacking statistical *a priori* signal information, the most popular next choice is to use the Maximum-Likelihood Method[3-5] but then no optimal properties can be guaranteed except, in some cases, minimax-type properties. The most ambitious attempt at using maximum *a posteriori* probability as a basis for receiver design (and for a situation in which the *a priori* signal statistics are largely under the control of the system designer) is probably carried out by Youla;[6] but again, no optimal property related to the quality of results to be obtained has been demonstrated.

It is the purpose of this note to demonstrate such an optimal property for a restricted class of receivers of the Youla type; namely, that the probability density function of the resulting linear error after demodulation is maximized for zero error. The meaning of this statement should become evident from the analytical discussion which follows below.

Let M and N be independent random variables representing "message" and "noise" respectively, with probability functions $P_M(M)$ and $P_N(N)$. We assume these functions to be differentiable. Let $R = F(M, N)$ be a given function of M and N which represents the input to the receiver.

We will assume that $(\partial F(M, N)/\partial N)$ is one-sided, and for simplicity we take $(\partial F(M, N)/\partial N)$. The inverse transformation, $N = \Phi(M, R)$, along with $(\partial^2\Phi/\partial M\,\partial R)$ are assumed to exist. From $N = \Phi(M, R)$ we have

$$1 = \frac{\partial \Phi}{\partial R}\frac{\partial F}{\partial N},$$

so that $(\partial\Phi(M, R)/\partial R) > 0$. First, we ask the following question: If R is given, what is the conditional probability that M lies in the interval $(M, M + dM)$? Let A be the event signifying that M lies in the interval $(M, M + dM)$, and let B denote the event signifying that R lies in the interval $(R, R + dR)$. We are interested in obtaining $P(A/B)$. From Bayes' theorem we have

$$P(A/B) = \frac{P_M(M)\,dMP(B/A)}{p_R(R)\,dR} \tag{1}$$

where $P(B/A)$ is the conditional probability that R lies in the interval $(R, R + dR)$, given M.

Now the probability that R exceeds R_0, given M, is

$$P[F(M, N) > R_0 \mid M] = P[N > \Phi(M, R_0)] = \int_{\Phi(M, R_0)}^{\infty} P_N(N)\,dN \tag{2}$$

so that

$$P(B/A) = \left[-\frac{\partial}{\partial R}\int_{\Phi(M,R)}^{\infty} P_N(N)\,dN\right]dR = P_N[\Phi(M, R)]\frac{\partial\Phi(M, R)}{\partial R}\,dR. \tag{3}$$

Thus (1) can be written

$$P(A/B) = P_M(M)P_N[\Phi(M, R)]\cdot\frac{\partial\Phi(M, R)}{\partial R}\Big/P_R(R). \tag{4}$$

If one chooses that estimate of M, for a given R, which maximizes (4), then the estimate of M is obtained by solving

$$\frac{\partial}{\partial M}\left\{P_M(M)P_N[\Phi(M, R)]\cdot\frac{\partial\Phi(M, R)}{\partial R}\right\} = 0 \tag{5}$$

for M in terms of R, for example, $M = f(R)$. Eq. (5) is the known maximum-likelihood method for estimating the message, M.

Now let us consider the following problem: Suppose that an estimate of M, the receiver output signal, is to be of the form $S = g(R) = g(F(M, N))$ for any given R. We assume that the inverse transformation $R = \phi(S)$ exists and is differentiable. The linear error in determining M is given by $\epsilon = S - M$, and this error will have a probability function given by $P_\epsilon(\epsilon)$. Every choice, $g(R)$, will yield a value of $P_\epsilon(0)$. We inquire as to the choice of $g(R)$ such that $P_\epsilon(0)$ shall be a maximum among all possible choices of the receiver function $g(R)$. We will show that the maximum *a posteriori* probability method discussed above will give just this result. Now,

$$P[F(M, N) > R_0] = \int_{-\infty}^{\infty}\int_{N=\Phi(M,R_0)}^{\infty} P_M(M)P_N(N)\,dN\,dM \tag{6}$$

so that

$$P_R(R) = \int_{-\infty}^{\infty} P_M(M)P_N[\Phi(M, R)]\cdot\frac{\partial\Phi(M, R)}{\partial R}\,dM. \tag{7}$$

Let A denote the probability of the event $(M, M + dM)$, and let C denote the probability of the event $(S, S + dS)$, with $S = g(R)$.
From

$$\begin{aligned} P(S > S_0 \mid M) &= P[R > \phi(S_0) \mid M] \\ &= P[F(M, N) > \phi(S_0) \mid M] \\ &= P\{N > \Phi[M, \phi(S_0)]\} \\ &= \int_{\Phi[M,\phi(S_0)]}^{\infty} P_N(N)\,dN \end{aligned} \tag{8}$$

we note that

$$P(C/A) = P_N\{\Phi[M, \phi(S)]\}\cdot\frac{\partial\Phi[M, \phi(S)]}{\partial\phi}\phi'(S)\,dS \tag{9}$$

so that Bayes' theorem yields

$$P(A/C) = P_M(M)P_N\{\Phi[M, \phi(S)]\}\cdot\frac{\partial\Phi[M, \phi(S)]}{\partial\phi}\,dM\Big/P_S(S). \tag{10}$$

* Received by the PGIT, July 15, 1959.

[1] J. L. Lawson and G. E. Uhlenbeck, "Threshold Signals," vol. 25, Radio Lab. Series, McGraw-Hill Book Co., Inc., pp. 167-173; 1950.

[2] D. Middleton, "Statistical theory of signal detection," IRE TRANS. ON INFORMATION THEORY, vol. IT-3, pp. 26-51; March, 1954.

[3] P. M. Woodward, "Probability and Information Theory with Applications to Radar," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 62-80; 1953.

[4] B. J. DuWaldt, "Inverse probability in angular tracking radars," IRE TRANS. ON INFORMATION THEORY, vol. IT-2, pp. 38-42; March, 1956.

[5] D. Slepian, "Estimation of signal parameters in the presence of noise," IRE TRANS. ON INFORMATION THEORY, vol. IT-3, pp. 68-89; March, 1954.

[6] D. C. Youla, "The use of the method of maximum likelihood in estimating continuous-modulated intelligence which has been corrupted by noise," IRE TRANS. ON INFORMATION THEORY, vol. PGIT-3, pp. 90-105; March, 1954.

rom

$$P(S - M > \epsilon)$$

$$= \int_{-\infty}^{\infty} \int_{S=\epsilon+M}^{\infty} P_s(S)P(A/C)\, dS\, dM \tag{11}$$

t follows that

$$P_\epsilon(\epsilon)$$

$$= \int_{-\infty}^{\infty} P_M(M)P_N\{\Phi[M, \phi(\epsilon + M)]\}$$

$$\frac{\partial \Phi[M, \phi(\epsilon + M)]}{\partial \phi}\, \phi'(\epsilon + M)\, dM$$

$$P_\epsilon(0)$$

$$= \int_{-\infty}^{\infty} P_M(M)P_N\{\Phi[M, \phi(M)]\}$$

$$\frac{\partial \Phi[M, \phi(M)]}{\partial \phi}\, \phi'(M)\, dM. \tag{12}$$

We wish to determine $\phi(M)$ which maximizes $P_\epsilon(0)$. One notes that the substitution $R = \phi(M)$ reduces the integrand of (12) to just that expression which is extremalized in (5), so that $P_\epsilon(0)$ is indeed a maximum for $M = f(R)$ as given by (5).

The curves in Fig. 1 illustrate our results. The curve (a) typifies a probability density function for the linear error for which the maximum-likelihood method is used for estimating the message, while curve (b) represents a p.d.f. for the linear error for an arbitrary continuous rule for determining the unknown message.

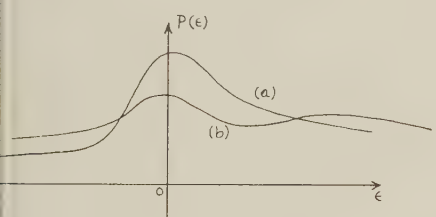


Fig. 1

An alternative proof of our result as suggested by the reviewer proceeds as follows. Let the random variables x and y denote "message" and "received signal," respectively. This communication process is characterized by three frequency functions, $f(x)$, $g(y)$, and $h(x/y)$, such that

1) $\int_A f(x)\, dx$ = the *a priori* probability that the message x is in A;
2) $\int_B g(y)\, dy$ = the probability of observing the received signal y in the set B; and
3) $\int_A h(x/y)\, dx$ = the *a posteriori* probability that the received signal y originated from a message x belonging to the set A.

Of necessity,

$$f(x) = \int_{-\infty}^{\infty} g(y)h(x/y)\, dy. \tag{13}$$

For any fixed $y = y_0$ we postulate the existence of a unique finite x_0 such that

$$h(x_0/y_0) \geqq h(x/y_0), \qquad -\infty < x < \infty. \tag{14}$$

This correspondence $y_0 \rightarrow x_0$ determines a single-valued function $x = m(y)$ such that $m(y)$ is the *a posteriori* maximum-likelihood estimate of x, given y.

Let $x = n(y)$ denote any arbitrary continuous receiver rule for determining the message x. From the definition of $m(y)$ it follows that

$$h(m(y)/y) \geqq h(n(y)/y). \tag{15}$$

$P_\epsilon(0)$ and $P_\epsilon{}^*(0)$ as given by (12) for the maximum-likelihood estimate of x and the arbitrary receiver rule for determining x can be expressed by

$$P_\epsilon(0) = \int_{-\infty}^{\infty} g(y)h(m(y)/y)\, dy; \tag{16}$$

$$P_\epsilon{}^*(0) = \int_{-\infty}^{\infty} g(y)h(n(y)/y)\, dy$$

As a consequence of (15) it follows that $P_\epsilon(0) \geqq P_\epsilon{}^*(0)$, Q.E.D.

Harry Lass
Jet Propulsion Lab.-CIT
Pasadena, Calif.
Robert M. Stewart
Litton Industries
Beverly Hills, Calif.

Two Notes on a Markoff Envelope Process*

In a recent paper[1] Pierce has shown that the envelope of the output of a high-Q singly tuned RLC filter is a first-order Markoff process when the filter input is white Gaussian noise. A simple proof of this follows from the fact that the transition probability density function (p.d.f.)

$$P(r_0 \mid r, t) = \frac{r}{1 - \mu^2} \cdot \exp\left[-\frac{r^2 + \mu^2 r_0^2}{2(1 - \mu^2)}\right] I_0\left(\frac{\mu r r_0}{1 - \mu^2}\right), \tag{1}$$

$$\mu = e^{-t},$$

satisfies a Fokker-Planck equation[2]

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial r}[A(r)P] + \tfrac{1}{2}\frac{\partial^2}{\partial r^2}[B(r)P] \tag{2}$$

with

$$A(r) = \frac{1}{r} - r, \qquad B(r) = 2. \tag{3}$$

$P(r_0 \mid r, t)\, dr$ is the probability that the envelope has a value between r and $r + dr$ at time t, given that $r = r_0$ at $t = 0$. (Here we have chosen units so that, in Pierce's notation, $\psi_0 = 1$, $k = 8$, $\alpha = 2$, $f_m \gg 1$.) The Fokker-Planck differential equation defines a first-order Markoff process and provides assurance that the transition p.d.f. $P(r_0 \mid r, t)$ obeys the Smoluchowski equation.[2]

(2) and (3) can be derived from the fact that the in-phase and quadrature components $x(t)$ and $y(t)$ of the output of the high-Q filter are independent first-order Gaussian Markoff processes with transition p.d.f.'s obeying[3] (2) with $A(r) = -r$, $B(r) = 2$. If we consider the diffusion of a particle in the XY-plane, starting at (x_0, y_0), its density function $p(x, y, t)$ at time t must satisfy the equation

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial}{\partial x}(xp) + \frac{\partial}{\partial y}(yp) = \frac{\partial p}{\partial t}, \qquad x \neq x_0, \qquad y \neq y_0, \tag{4}$$

which is obtained by combining the above-mentioned Fokker-Planck equations for the processes $x(t)$ and $y(t)$. Changing to polar co-ordinates $x = r\cos\theta$, $y = r\sin\theta$ in (4), we get

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 p}{\partial \theta^2} + \frac{\partial}{\partial r}(rp) + p = \frac{\partial p}{\partial t}. \tag{5}$$

The envelope process $r(t)$ is just the radius vector to the diffusing particle. Since the initial phase θ_0 is uniformly distributed in $0 \leq \theta_0 < 2\pi$, the density function $P(r_0 \mid r, t)$ for the envelope is independent of the angle θ. This function represents the density of particles in an annulus of area $2\pi r dr$ at radius r in the XY-plane, and must therefore be $2\pi r$ times the function p appearing in (5). Making the substitution $P(r_0 \mid r, t) = 2\pi r p(x, y, t)$ in (5), we obtain (2), (3) immediately.

The Fokker-Planck equation (2) can be used to find the p.d.f. $Q(r_0 \mid a, t)$ of the first time t that the envelope $r(t)$ crosses

* Received by the PGIT, June 24, 1959; March 23, 1959.

[1] J. N. Pierce, "A Markoff envelope Process," IRE Trans. on Information Theory, vol. IT-4, pp. 163–166; December, 1958.

[2] M. C. Wang and G. E. Uhlenbeck, "On the theory of the Brownian motion II," *Rev. Mod. Phys.*, vol. 17, pp. 323–342; April, July, 1945. (Reprinted in N. Wax, "Noise and Stochastic Processes," Dover Publications, Inc., New York, N. Y., pp. 113–132; 1954.

[3] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion," *Phys. Rev.*, vol. 36, pp. 823–841; Sept. 1, 1930. (Reprinted in N. Wax, *op. cit.*, pp. 93–111.)

a threshold level $r = a$, given that it had the value r_0 at $t = 0$. This first-passage-time problem can be solved by the methods of Siegert,[4] for which we need the Laplace transform

$$P_\lambda(r_0 \mid r) = \int_0^\infty e^{-\lambda t} P(r_0 \mid r, t)\, dt \quad (6)$$

of the transition p.d.f. It can be found[5] by solving the Laplace-transformed Fokker-Planck equation (2) with appropriate boundary conditions, and it turns out to be

$$\begin{aligned} P_\lambda(r_0 \mid r) &= \frac{r}{2}\,\Gamma(\lambda/2)e^{-r^2/2}\Psi(\lambda/2, 1; r_0^2/2) \\ &\quad \cdot \Phi(\lambda/2, 1; r^2/2), \qquad r < r_0 \\ &= \frac{r}{2}\,\Gamma(\lambda/2)e^{-r^2/2}\Phi(\lambda/2, 1; r_0^2/2) \\ &\quad \cdot \Psi(\lambda/2, 1; r^2/2), \qquad r > r_0, \end{aligned} \quad (7)$$

where Φ and Ψ are the confluent hypergeometric functions as defined by Erdélyi *et al.*[6] According to Siegert,[4] the Laplace transform $Q_\lambda(r_0 \mid a)$ of the first-passage-time p.d.f. $Q(r_0 \mid a, t)$ is then

$$Q_\lambda(r_0 \mid a) = \frac{\Phi(\lambda/2, 1; r_0^2/2)}{\Phi(\lambda/2, 1; a^2/2)}. \quad (8)$$

It must be inverted to find the p.d.f. itself. By the usual techniques one can obtain a series of the form

$$Q(r_0 \mid a, t) = \sum_k A_k e^{\lambda_k t}, \quad (9)$$

where the A_k are the residues of $Q_\lambda(r_0 \mid a)$ at the roots λ_k of

$$\Phi(\lambda_k/2, 1; a^2/2) = 0. \quad (10)$$

These roots can be calculated approximately[5] by interpolating between the zeros of the Laguerre polynomials $L_m(x)$, since

$$\Phi(-m, 1; x) = \Gamma(m + 1)L_m(x),$$

$$x = a^2/2, \quad m = \text{positive integer}.$$

Those zeros have been tabulated by Salzer and Zucker.[7] Some curves of the first-passage-time p.d.f. for particular values of a and for an envelope r initially Rayleigh distributed in $0 < r_0 < a$ are given in the report cited.[5]

[4] A. J. F. Siegert, "On the first passage time probability function," *Phys. Rev.*, vol. 81, pp. 617–623; February 15, 1951.

[5] C. W. Helstrom, "The accuracy of probability distributions computed by the sampling approximation," Westinghouse Res. Rep. No. 8-1259-R2; May 21, 1956.

[6] A. Erdélyi *et al.*, "Higher Transcendental Functions," McGraw-Hill Book Co., Inc., New York, N. Y., 1953. See ch. 6, pp. 248–295.

[7] H. E. Salzer and R. Zucker, "Tables of the zeros and weight factors of the first fifteen Laguerre polynomials," *Bull. Am. Math. Soc.*, vol. 55, pp. 1004–1012; October, 1949. (Reprinted in "Tables of Functions and of Zeros of Functions," Nat'l Bur. of Standards, Appl. Math. Series No. 37, Washington, D. C., pp. 191–199; 1954.)

In particular one can easily calculate the mean first-passage-time $\bar{t}\,(r_0 \mid a)$ for an envelope starting at r_0 and crossing a level $a > r_0$, by the formula[5]

$$\begin{aligned} \bar{t}(r_0 \mid a) &= -\frac{\partial}{\partial\lambda} Q_\lambda(r_0 \mid a)\,|_{\lambda=0} \\ &= \frac{1}{2}\left[\frac{\partial}{\partial\lambda}\,\Phi(\lambda, 1; a^2/2)\right. \\ &\quad \left. - \frac{\partial}{\partial\lambda}\,\Phi(\lambda, 1; r_0^2/2)\right]\Big|_{\lambda=0} \\ &= \frac{1}{2}\int_{r_0^2/2}^{a^2/2} \frac{e^s - 1}{s}\, ds \\ &= \tfrac{1}{2}[\bar{E}i(a^2/2) - \bar{E}i(r_0^2/2) \\ &\quad - 2 \ln (a/r_0)], \end{aligned} \quad (11)$$

where the function $\bar{E}i(x)$ is tabulated in Jahnke-Emde.[8] The above expression is most easily derived by writing out the hypergeometric series for each term, differentiating it with respect to λ, and setting λ equal to zero. The resulting series is then the same as is obtained when the above integrand is expanded in a power series in s and integrated term by term.

Carl W. Helstrom
Dept. of Math.
Westinghouse Res. Lab.
Pittsburgh, Pa.

[8] E. Jahnke and F. Emde, *Tables of Functions*, Dover Publications, Inc., New York, N. Y., pp. 6–8; 1945.

In a recent article by Pierce,[1] a special case for a Markoff envelope process was discussed. While the referenced paper would serve admirably as a tutorial exposition of the manipulative mathematics associated with the elementary calculus of probability theory, the conclusions reached by Pierce can be arrived at in a much more direct and simple fashion. In fact, it is possible to draw more general conclusions regarding the order properties of envelope processes.

The effects of a linear network whose transfer function has a finite number of poles on the dimensionality of a random process have been discussed by several authors.[9,10] It was discussed in a particularly straightforward and concise way by Gold and Young,[2] who demonstrated that a linear network whose transfer characteristics are determined by a constant coefficient linear differential equation of order N, introduces an added dimensionality of N to any statistical process which is passed through it. This dimensionality property is not restricted to Gaussian processes; Heilfron,[3] in fact, applies this property to cases where general, zero-memory nonlinearities exist.

The linear network which Pierce[1] has selected possesses two poles (a conjugate pair). Since the envelope detector is a zero memory device, the dimensionality of the process at its output cannot, by definition, exceed the dimensionality of the process at its input. Hence, at most, the statistics of the envelope detector output are specified by a third order process or second order Markoff process. Pierce's assumption of a network having, as a close approximation, "narrow band" symmetry about a nominal center frequency permits, in the manner of Rice,[11] the resolution of the Gaussian noise into in-phase and quadrature components about the assumed center frequency. As is well known, these components also have Gaussian distributions and, within the validity that symmetry exists, are independent processes. As a result of the assumed narrow-band symmetry, the equivalent network transfer function effective with each component possesses but one simple pole. This is all tacitly present in Pierce's development, though he did not explicitly point out this simple but significant property. Thus, for an input of white Gaussian noise, both the in-phase and quadrature components are defined respectively by independent first-order Markoff processes. Demonstration that the envelope process also becomes a first-order Markoff process then becomes trivial.

These conclusions can easily be extended to a more general case. If the predetection linear network is such that its transfer function is rational, having $2N$ poles, then, at most, the envelope function is a vector of a $(2N)$th order Markoff process. In particular, if a specification of narrow-band symmetry may be validly applied, the resulting envelope function becomes a vector of an Nth-order Markoff process.

C. T. Isley
Communications Div.
Hughes Aircraft Co.
Los Angeles, Calif.

[9] B. Gold and G. O. Young, "The response of linear systems to non-Gaussian noise," IRE Trans. on Information Theory, vol. IT-3, pp. 63–67; March 1954.

[10] J. Heilfron, "On the response of a certain class of systems to random inputs," IRE Trans. on Information Theory, vol. IT-1, pp. 59–61; March, 1955.

[11] S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 24, pp. 75–79; January, 1945.

A Note on Angle Modulation by a Mixture of a Periodic Function and Noise*

This paper examines some properties of the waveform

* Received by the PGIT, April 13, 1959.

$$Y(t) = \cos [\omega t + P(t) + N(t)].$$

This expression represents angle-modulation of a carrier ω by a mixture of a periodic function $P(t)$ and a noise function $N(t)$. A result formerly stated, to the effect that under certain conditions $Y(t)$ can be approximately represented as a sum of waveforms, each of which is the result of $N(t)$ modulating one of the sinusoidal components of $\cos [\omega t + P(t)]$, is reviewed. This approximate result is supplemented by exact statements concerning the correlation function of $Y(t)$; one of these statements is, that under certain easily satisfied conditions, the correlation function of $Y(t)$ is that of $\cos [\omega t + P(t)]$ multiplied by a factor depending only on the properties of $N(t)$.

Introduction

This note deals with the following question: if one is attempting to produce a waveform having a line spectrum, by angle-modulating a carrier by a periodic function, to what extent will the unavoidable addition of noise to the nominally-periodic modulating function cause the resulting waveform to lose its line spectrum character? Some of this noise may be considered as deriving from the carrier fluctuations. In a mathematical sense it is, of course, true that any amount of a non-periodic function or noise will destroy the line spectrum. However, the nontrivial problem is really the speed at which spaces between the lines in the spectrum fill in, as the power or other properties of the noise change in a manner relative to those of the periodic modulating function or the carrier. A corresponding question may be asked concerning the autocorrelation function.

An Approximate Relation in the Time Domain

A result of some interest in this connection is presented as relation (14) below. Let

$$y(t) = \cos [\omega t + P(t)] \equiv \cos [\phi(t)] \quad (1)$$

represent the phase modulation of a carrier ω by a periodic function $P(t)$ having a basic period $T = 1/a$. It can easily be shown that $y(t)$ has a trigonometric series expansion of the form,

$$y(t) = \sum_n d_n \cos [(\omega + 2\pi na)t + \beta_n], \quad (2)$$

which represents a line spectrum.

If we add a noise function $N(t)$[1] to the modulating function $P(t)$ we have,

$$Y(t) = \cos [\omega_0 t + P(t) + N(t)] = \cos [\phi(t) + N(t)] \quad (3)$$

representing the resulting waveform $Y(t)$.

[1] Throughout this paper it is assumed that the random function $N(t)$ has appropriate properties of stationarity and ergodicity.

Go through the following:

$$Y(t) = \cos \phi(t) \cos N(t) - \sin \phi(t) \sin N(t), \quad (4)$$

$$\sin \phi(t) = \cos \left(\phi(t) - \frac{\pi}{2}\right) = \cos \left[\omega t - \frac{\pi}{2} + P(t)\right] = \cos \left[\omega\left(t - \frac{\pi}{2\omega}\right) + P(t)\right] \quad (5)$$

$$= \cos \left[\omega\left(t - \frac{\pi}{2\omega}\right) + P\left(t - \frac{\pi}{2\omega}\right) + P(t) - P\left(t - \frac{\pi}{2\omega}\right)\right] \quad (6)$$

$$= \cos \left[\phi\left(t - \frac{\pi}{2\omega}\right) + P(t) - P\left(t - \frac{\pi}{2\omega}\right)\right]. \quad (7)$$

If on the average, $P(t) - P(t - \pi/2\omega)$ is very small (that is, if it is a small part of a radian), or if

$$P(t) - P\left(t - \frac{\pi}{2\omega}\right) \ll 1, \quad (8)$$

then (7) can be replaced by

$$\sin \phi(t) \approx \cos \left[\phi\left(t - \frac{\pi}{2\omega}\right)\right] \equiv y\left(t - \frac{\pi}{2\omega}\right). \quad (9)$$

The restriction on $P(t) - P(t - \pi/2\omega)$ means that the modulating phase changes very little during the time $\pi/2\omega$, a quarter of a carrier cycle. Clearly, this is a case of practical interest. With (8), then, we use (2) and write

$$y\left(t - \frac{\pi}{2\omega}\right) = \sum_n d_n \cos \left[(\omega + 2\pi na)\left(t - \frac{\pi}{2\omega}\right) + \beta_n\right] \quad (10)$$

$$= \sum_n d_n \cos \left[\omega t + 2\pi nat - \frac{\pi}{2} - \frac{2\pi na\pi}{2\omega} + \beta_n\right] = \sum_n d_n \sin \left[\omega t + 2\pi nat + \beta_n - \frac{2\pi na\pi}{2\omega}\right]. \quad (11)$$

If the quantities $(2\pi na\pi/2\omega)$ are (for any values of n which are of interest) small parts of a radian, or if

$$(2\pi na\pi/2\omega) \ll 1, \quad (12)$$

(in other words if the effective bandwidth of the waveform (1) is small compared to the carrier frequency), we may write from (11)

$$\sin \phi(t) \approx y\left(t - \frac{\pi}{2\omega}\right) \approx \sum_n d_n [\sin \omega t + 2\pi nat + \beta_n]. \quad (13)$$

The latter approximation is not consistent with the former.

Inserting (13) in (4) we have

$$Y(t) \approx \cos N(t) \sum d_n \cdot \cos [\omega t + 2\pi nat + \beta_n] - \sin N(t) \cdot \sum d_n \sin [\omega t + 2\pi nat + \beta_n] \approx \sum_n d_n \cos [\omega t + 2\pi nat + \beta_n + N(t)]. \quad (14)$$

The last result may be stated in the following manner: phase modulation with $N(t)$ of the complex waveform (1) or (2) has the same effect as if $N(t)$ were used to modulate each of the *components* of the waveform (2). Note that we have used the special conditions (8) and (12) to establish the approximate equality (14).[2]

The Correlation Function

The approximate result (14) refers to the time function $Y(t)$. We now derive some exact results which refer to the autocorrelation function of $Y(t)$, $R_Y(\tau)$. We consider (3) for $Y(t)$. Since $Y(t)$ has a non-random part in the argument of the cosine, the statistical (ensemble) properties of the product $Y(t)Y(t - \tau)$ depend upon t. However, we can still get a useful correlation function by averaging over t. For our purposes, therefore, we shall define $R_Y(\tau)$ as

$$R_Y(\tau) = \langle \overline{Y(t)\, Y(t - \tau)} \rangle, \quad (15)$$

where the bar denotes time average and the brackets denote ensemble average. This corresponds to the common experimental situation in which (in one idealized model) one takes a time average of a correlation product and then proceeds to consider the ensemble average or expected value.

[2] N. D. Blackman, in Tech. Mem. EDL-M43 of the Electronic Def. Lab., April 20, 1955, gives the result (14) but invokes only the aid of (12). It seems to the writer that (8) must also be assumed. Eq. (12) does not imply (8); of course, neither of the treatments tells us how accurate (14) is.

Consider, then, the time average

$$\overline{\cos\,[\phi(t) + N(t)]\cos\,[N(t-\tau) + N(t-\tau)]}$$
$$= \tfrac{1}{2}\,\overline{\cos\,[\phi(t) - \phi(t-\tau) + N(t) - N(t-\tau)]}$$
$$+ \tfrac{1}{2}\,\overline{\cos\,[\phi(t) + \phi(t-\tau) + N(t) + N(t-\tau)]}. \quad (16)$$

The last line may be replaced by

$$\tfrac{1}{2}\,\overline{\cos\,[N(t) + N(t-\tau)]\cos\,[\phi(t) + \phi(t-\tau)]}$$
$$-\tfrac{1}{2}\,\overline{\sin\,[N(t) + N(t-\tau)]\sin\,[\phi(t) + \phi(t+\tau)]}. \quad (17)$$

Upon taking ensemble averages, (17) will be

$$\tfrac{1}{2}\langle\cos\,[N(t) + N(t-\tau]\rangle\,\overline{\cos\,[\phi(t) + \phi(t+\tau)]}$$
$$-\tfrac{1}{2}\langle\sin\,[N(t) + N(t-\tau)]\rangle\,\overline{\sin\,[\phi(t) + \phi(t+\tau)]}. \quad (18)$$

From (1), $\phi(t)$ is of the form $\omega t + P(t)$. Thus, the time averages appearing in (18) of the form

$$\overline{\cos\,[2\omega t - \omega\tau + P(t) - P(t-\tau)]}, \quad (19a)$$

and

$$\overline{\sin\,[2\omega t - \omega\tau + P(t) - P(t-\tau)]}. \quad (19b)$$

Now $P(t) - P(t-\tau)$ has the same period as does $P(t)$, namely, T. It can be proven[3] that (19a) and (19b) approach zero if the average is taken during a sufficiently long time compared with T and if

$$T \neq \frac{2\pi n}{\omega}, \quad (20)$$

where n is an integer. In this sense, then, (19a) and (19b) and (17) may be replaced by zero, and the ensemble average of (16) becomes

$$R_Y(\tau) = \tfrac{1}{2}\langle\cos\,[\phi(t) - \phi(t-\tau) + N(t) - N(t-\tau)]\rangle. \quad (21)$$

Expanding (21), we have for our correlation function,

$$R_Y(\tau) = \tfrac{1}{2}\langle\cos\,[N(t) - N(t-\tau)]\rangle\,\overline{\cdot\cos\,[\phi(t) - \phi(t-\tau)]}$$
$$= \tfrac{1}{2}\langle\sin\,[N(t) - N(t-\tau)]\rangle\,\overline{\cdot\sin\,[\phi(t) - \phi(t-\tau)]} \quad (22)$$

or,

$$R_Y(\tau) = \overline{\langle Y(t)\,Y(t-\tau)\rangle}$$
$$= C_N(\tau)C_Y(\tau) - S_N(\tau)S_Y(\tau), \quad (23)$$

[3] To be demonstrated in the appendix.

where

$$C_N(\tau) = \langle\cos\,[N(t) - N(t-\tau)]\rangle,$$
$$S_N(\tau) = \langle\sin\,[N(t) - N(t-\tau)]\rangle,$$
$$C_Y(\tau) = \tfrac{1}{2}\,\overline{\cos\,[\phi(t) - \phi(t-\tau)]},$$
$$S_Y(\tau) = \tfrac{1}{2}\,\overline{\sin\,[\phi(t) - \phi(t-\tau)]}. \quad (24)$$

It may be of some interest to write (23) in another form. Define the complex correlation coefficients

$$r_N(\tau) = \langle f_N(t)f_N^*(t-\tau)\rangle \quad (25)$$

and

$$r_y(\tau) = \overline{f_y(t)f^*{}_y(t-\tau)},$$

where

$$f_N(t) = e^{iN(t)} \quad (26)$$

and

$$f_y(t) = e^{i\phi(t)}.$$

Then

$$R_Y(\tau) = \text{Real part of } [r_N(\tau)r_y(\tau)]. \quad (27)$$

Returning to the basic results (22) or (23), we point out that if $N(t)$ were identically zero for all (t), (22) would reduce to

$$\tfrac{1}{2}\,\overline{\cos\,[\phi(t) - \phi(t-\tau)]} \equiv C_y(\tau). \quad (28)$$

Therefore, (23) must be the same as

$$\overline{\cos\,\phi(t)\,\cos\,\phi(t-\tau)}; \quad (28a)$$

i.e., the correlation function of $y(t) = \cos\,[\phi(t)]$. To show this directly, we may replace (28a) by

$$\tfrac{1}{2}\,\overline{\cos\,[\phi(t) - \phi(t-\tau)]} + \tfrac{1}{2}\,\overline{\cos\,[\phi(t) + \phi(t-\tau)]},$$

whose second term is one-half of (19a) or zero, which demonstrates the equality mentioned.

The result (22) or (23) may now be expressed in the following way: when $N(t)$ is added to $\phi(t)$, the correlation function $R_Y(\tau)$ of the resulting noise-modulated function $Y(t) \equiv \cos\,[\phi(t) + N(t)]$, is the same as that of $\cos\,\phi(t)$, namely, $C_y(\tau)$, multiplied by $C_N(\tau)$, plus a "correction term" $-S_N(\tau)S_y(\tau)$. [If (19a) and (19b) were not zero, which could happen—for example, if (20) were not satisfied—$C_y(\tau)$ would not be identical with the correlation function of $\cos\,\phi(t)$; furthermore, the expression for $R_y(\tau)$ would contain the two additional terms exhibited in (18).]

The "correction term" is zero in important cases. Thus, if $N(t) - N(t-\tau)$ has a probability density function which is symmetrical about zero, the average of $\sin\,[N(t) - N(t-\tau)]$, or $S_N(\tau)$, is zero, and this causes the correction term to be zero. One important case in which this symmetry exists is that in which $N(t)$ is zero-centered Gaussian noise. Another is that in which $N(t)$ is the result of passing Gaussian noise through a symmetrical limiter, or a nonlinear circuit operating on positive and negative amplitudes in a symmetrical manner.

With the correction term zero, our result becomes

$$R_Y(\tau) = \tfrac{1}{2}\langle\cos\,[N(t) - N(t-\tau)]\rangle\,\overline{\cos\,[\phi(t) - \phi(t-\tau)]}$$
$$= \langle\cos\,[N(t) - N(t-\tau)]\rangle\,\overline{\cos\,\phi(t)\,\cos\,\phi(t-\tau)}$$
$$= C_N(\tau)C_y(\tau), \quad (29)$$

where $C_y(\tau)$ is the correlation function of $y(t)$. The quantity $C_N(\tau)$ can apparently be evaluated in a number of cases. In particular, one may compute it in straightforward manner, in the Gaussian case, and obtain,

$$C_N(\tau) = \langle\cos\,[N(t) - N(t-\tau)]\rangle$$
$$= e^{-\sigma^2}[1 - \rho(\tau)], \quad (30)$$

where σ^2 is the mean square value of $N(t)$; *i.e.*,

$$\sigma^2 = \langle N^2(t)\rangle$$

and $\sigma^2\rho(\tau)$ is the correlation function of $N(t)$; *i.e.*,

$$\sigma^2\rho(\tau) = \langle N(t)N(t-\tau)\rangle.$$

Summary

Thus, the approximate relation (14) may be augmented by exact statements, (29) and (22), or (23). Statement (29) says that all functions $y(t)$ [obeying (20)] will be affected in the same way with respect to the correlation function by the addition of the modulating phase $N(t)$. That is, that a simple sinusoidal function $y(t) + \cos\,\omega t$ will have its correlation function multiplied by $C_N(\tau)$ as would a more complicated function $y(t) = \cos\,[\omega t + P(t)]$. Furthermore, since the correlation function of $y(t)$ is the sum of the correlation functions of the individual sinusoidal components of $y(t)$, the effect of $N(t)$ will be the production of a correlation function $R_Y(\tau)$ which is the sum of the individual correlation functions of $y(t)$, each multiplied by $C_N(\tau)$. This can now be related to (14), which is the superposition of sinusoidal waves each modulated with the same noise $N(t)$. One

can see that the correlation function of (14) would be such a sum of correlation functions. The derivation of (14), however, is contingent upon certain assumptions for the depth of modulation or resultant spectra, whereas (29) is not dependent upon such assumptions.

Since the factor $C_N(\tau) = \langle \cos [N(t) - N(t - \tau)] \rangle$ generally represents (as seen in the Gaussian case), a damping effect on the correlation function considered as a function of τ this damping effect will be *the same whether $y(t)$ is a simple sinusoidal wave or a more complex wave form.*

The effect of $N(t)$ upon the "power spectrum" or spectral density function of the wave form (1) (which is a series of delta functions) is the convolving of each of the delta functions with the spectrum associated with the factor $C_N(\tau)$. That is, that the new spectrum [Fourier transform of (29)] will be the superposition of individual spectra, centered at the frequencies $\omega/2\pi \pm n/T$ proportional to the strength of the original delta functions there.

Thus, the effect of $N(t)$ is to "smear" each spectral component of the original function (1) in the same manner.

Appendix

We wish to prove that

$$\lim_{x\to\infty} F(x) = 0,$$

where

$$F(x) = \frac{1}{x}\int_0^x \cos [\omega t + P(t)]\, dt, \quad (31)$$

and where $P(t)$ is periodic.

There is no loss of generality if we let the period of $P(t)$ be 2π. Then consider the quantity $F(2m\pi)$ where m is an integer:

$$F(2m\pi) = \frac{1}{2m\pi}\int^{m2\pi} \cos [\omega t + P(t)]\, dt \quad (32)$$

$$\frac{1}{2m\pi}\int_0^{m2\pi} [e^{i(\omega t + Pt)} + c.c.]\, dt, \quad (33)$$

where "c.c." means "complex conjugate."

Since $P(t)$ is periodic with period 2π, $e^{iP(t)}$ also is. Let now $e^{iP(t)} = f(t)$ and consider the integral

$$\phi(2m\pi) = \frac{1}{2m\pi}\int_0^{2m\pi} f(t)e^{i\omega t}\, dt.$$

We have

$$\phi(2m\pi) = \frac{1}{2m\pi}\left\{\int_0^{2\pi} + \int_{2\pi}^{4\pi} + \cdots + \int_{(m-1)2\pi}^{m2\pi} f(t)e^{i\omega t}\, dt\right\}$$

$$= \frac{1}{m}\left\{1 + e^{i\omega 2\pi} + \cdots + e^{i\omega(m-1)2\pi}\right\} \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{i\omega t}\, dt$$

$$= \frac{1}{m}\,\frac{1 - e^{i\omega m 2\pi}}{1 - e^{i\omega 2\pi}} \cdot \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{i\omega t}\, dt$$

$$= \frac{1}{m}\,\frac{1 - e^{i\omega m 2\pi}}{1 - e^{i\omega 2\pi}}\,\phi(2\pi).$$

Since the maximum possible value of $1 - e^{i\omega m 2\pi}$ is 2, we have

$$\lim_{m\to\infty} \phi(m2\pi) = 0,$$

if $\phi(2\pi)$ is finite and if $1 - e^{i\omega\pi}$ is not zero. The latter condition will be satisfied if $\omega 2\pi \neq n2\pi$ or

$$2\pi \neq \frac{n2\pi}{\omega},$$

where n is an integer.

If the period of $P(t)$ were T, this condition would be replaced by

$$T \neq \frac{n2\pi}{\omega}.$$

The complex conjugate of $\phi(m2\pi)$ will have similar properties, and its limit will also be zero. Therefore, since

$$\lim_{x\to\infty} F(x) = \lim_{m\to\infty} F(m2\pi) + c.c.,$$

we have

$$\lim_{x\to\infty} F(x) = 0.$$

It will be noted that the only condition placed on $f(t)$ beside the periodicity already mentioned, is that the integral $\phi(x)$ exists for all values of x in the range $0 - 2\pi$. This is a light restriction and does not require $f(t)$ to be continuous. Thus, we have proven the fact that the average of $\cos [\omega t + P(t)]$ is zero, if $T \neq n2\pi/\omega$ with no other important restriction on $P(t)$. The basis of the present proof is taken from Bohr's "Almost Periodic Functions," page 51.[4]

Acknowledgment

The author wishes to thank M. Knopp for discussing this question and other possible proofs.

P. R. Karr
Ramo-Wooldridge
Div. of Thompson Ramo-Wooldridge, Inc.
Los Angeles, Calif.

[4] The discussion in Bohr's book concerns the case of continuous functions, but the result is seen to be valid for a wider class of functions.

Contributors

William M. Cowan was born in Chicago, Ill., on July 28, 1934. He received the B.S. degree in electrical engineering from Northwestern University, Evanston, Ill., in 1957, and the M.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1959. He held a Whitney Fellowship the first year at M.I.T., and worked as a teaching assistant during his last term. The title of his master's thesis was "Experimental Determination of Optimum Non-Linear Filters."

W. M. Cowan, Jr.

He was with Sylvania's Applied Research Laboratory during the summer of 1958, working on false alarm probabilities in a bank of parallel filters and on the estimation of doppler frequencies. He spent the summer of 1957 at Motorola, where he worked on the transistorization of the high frequency stages of their Handy Talky receiver. He also spent six quarters at Cook Research Laboratories while he was a co-op student at Northwestern. He rejoined Sylvania's Applied Research Laboratories, Waltham, Mass., as a permanent employee in February, 1959.

Mr. Cowan is a member of Tau Beta Pi, Eta Kappa Nu, and Pi Mu Epsilon.

❖

Bernard Friedland (S '52—A '55—M '58) was born in Brooklyn, N. Y., on May 25, 1930. He received his education at Columbia University where he was awarded the A.B. Degree in 1952, the B.S. degree in 1953, the M.S. degree in 1954 under a National Science Foundation Fellowship, and the Ph.D. degree in 1957.

B. Friedland

In 1954 he joined the Department of Electrical Engineering of Columbia as an instructor and became an assistant professor of electrical engineering in that department in 1957. He has taught courses in linear system theory, network theory, and sampled-data control systems, and is at present engaged in research in the areas of automatic control and sequential systems.

Dr. Friedland is a member of Phi Beta Kappa, Sigma Xi, and Tau Beta Pi.

Janis Galejs (A'52) was born in Riga, Latvia, on July 21, 1923. He received the Engineering Diploma in electrical engineering from the Technical University, Brunswick, Germany, in 1950, and the M.S. and Ph.D. degrees in electrical engineering from the Illinois Institute of Technology, Chicago, Ill., in 1953 and 1957, respectively.

J. Galejs

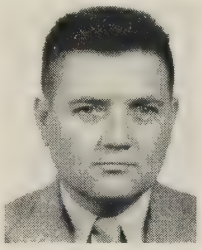
While attending I.I.T. he worked for the Cook Research Laboratory on fire control problems, radar, and communication systems.

He joined the Applied Research Laboratory of Sylvania Electric Products, Inc., in Waltham, Mass. in 1957 and is now studying radar systems.

Dr. Galejs is a member of Sigma Xi and Tau Beta Pi.

❖

E. T. Jaynes (SM '54) was born in Waterloo, Iowa, on July 5, 1922. He attended Cornell College and Iowa State University, receiving the B.A. degree in physics from the latter in 1942. He studied in the graduate school of the University of California in Berkeley and at Princeton University from which he received the M.A. degree in 1948 and the Ph.D. degree in theoretical physics in 1950.



E. T. Jaynes

From 1942 to 1946, he was engaged in microwave research and development as a project engineer at the Sperry Gyroscope Co., Garden City, N. Y., and in the combined research group of the Naval Research Laboratory.

Since 1950, he has been on the faculty of Stanford University, Stanford, Calif., and at present holds the titles of associate professor in the microwave laboratory and lecturer in physics.

❖

Max V. Mathews (S '53—A '55) was born in Columbus, Nebr., on November 13, 1926. He received a B.S. degree in electrical engineering from California Institute of Technology, Pasadena, Calif., in 1950, and M.S. and Sc.D. degrees in electrical engineering from Massachusetts Institute of Technology, Cambridge, Mass., in 1952 and 1954 respectively.

He joined the staff of the Bell Telephone Laboratories in 1955 where he is working in visual and acoustics research. His main activities have concerned the study of speech coding and recognition methods by means of digital computer simulation.

M. V. Mathews

Mr. Mathews is a member of the Acoustical Society of America, and Sigma Xi.

❖

Irving S. Reed was born in Seattle, Wash. in 1923. He received his Ph.D. degree in mathematics from the California Institute of Technology, Pasadena, Calif., in 1949.

I. S. Reed

He has been associated with the Lincoln Laboratory of the Massachusetts Institute of Technology, Lexington, Mass., for the past eight years. His interests are in mathematics, the design of computing machines, stochastic processes, and information theory.

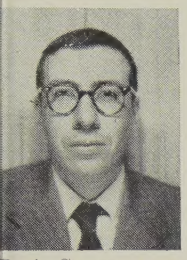
❖

Harold S. Shapiro was born on April 2, 1928, in Brooklyn, N. Y. He received the B.S. degree at City College, N. Y., in 1949, and the Ph.D. degree in mathematics at the Massachusetts Institute of Technology, Cambridge, Mass., in 1952.

H. S. Shapiro

From 1952–1954 he was a member of the technical staff of the Bell Telephone Laboratories at Murray Hill, N. J. where he worked on switching problems and prediction theory. In 1954 he joined the Institute of Mathematical Sciences at New York University as a research associate; he is now an assistant professor of mathematics. His main work has been in the theory of functions.

Richard A. Silverman (M '54—SM '58) as born on June 29, 1926, in Boston, Iass. He received the A.B. degree from Harvard University in 1946, the M.A. degree from Columbia University in 1948, and the Ph.D. degree from Harvard in 1951.



R. A. Silverman

For three years Dr. Silverman was associated with the Massachusetts Institute of Technology, first as a staff member of the Lincoln aboratory and then as a research associate the Department of Electrical Engieering.

Dr. Silverman is currently a research associate at the New York University Institute of Mathematical Sciences in the Division of Electromagnetic Research.

He is a member of Phi Beta Kappa, Sigma Xi, and the American Physical Society.

❖

Thomas E. Stern (S '54—M '57) was born in New Rochelle, N. Y., on March 29, 1930. He pursued his engineering education at the Massachusetts Institute of Technology, receiving his undergraduate education under the cooperative program in Electrical Engineering. After receiving the B.S. and M.S. degrees in 1953, he became a research assistant at the MIT Research Laboratory of Electronics, receiving the Sc.D. degree from MIT in 1956.

T. E. Stern

At present he is Assistant Professor of Electrical Engineering at Columbia University. His areas of research include analog computation, nonlinear network theory and information theory.

Professor Stern is a member of Eta Kappa Nu, and Sigma Xi.

INFORMATION FOR AUTHORS

Authors are requested to submit editorial correspondence or technical manuscripts to the Publications Chairman for possible publication in the PGIT TRANSACTIONS. Papers submitted should include a statement as to whether the material has been copyrighted, previously published, or accepted for publication elsewhere.

Papers should be written concisely, keeping to a minimum all introductory and historical material. It is seldom necessary to reproduce in their entirety previously published derivations, where a statement of results, with adequate references, will suffice.

To expedite reviewing procedures, it is requested that authors submit the original and two legible copies of all written and illustrative material. The manuscript should be double-spaced, and the illustrations drawn in India ink on drawing paper or drafting cloth. Each paper should include a carefully written abstract of not more than 200 words. Upon acceptance, papers should be prepared for publication in a manner similar to those intended for the PROCEEDINGS OF THE IRE. Further instructions may be obtained from the Publications Chairman. Material not accepted for publication will be returned.

IRE TRANSACTIONS ON INFORMATION THEORY is published four times a year, in March, June, September, and December. A minimum of one month must be allowed for review and correction of all accepted manuscripts. In addition, a period of approximately two months is required for the mechanical phases of publication and printing. Therefore, all manuscripts must be submitted three months prior to the respective publication dates.

All technical manuscripts and editorial correspondence should be addressed to George A. Deschamps, University of Illinois, Urbana, Ill. Local Chapter activities and announcements, as well as other nontechnical news items, should be addressed to Laurin G. Fischer, ITT Laboratories, 492 River Road, Nutley 10, N. J.

INSTITUTIONAL LISTINGS

The IRE Professional Group on Information Theory is grateful for the assistance given by the firms listed below and invites application for Institutional Listing from other firms interested in the field of Information Theory.

IBM RESEARCH, INTERNATIONAL BUSINESS MACHINES CORP., Yorktown Heights, N. Y.
Error Correcting & Detecting Codes, Theory of Assemblies & Automata, Information Networks, Reliability

REPUBLIC AVIATION CORP., Farmingdale, N. Y.
Aircraft, Missiles, Drones, Electronic Analyzers; U. S. Distr. of Alouette Turbine-Powered Helicopter